Chapter 0

Introduction

0.1 We will study geometry as it is used in physics.

0.2 Mathematics is the most exact possible thought.

0.3 Physics is the most exact description of nature; the laws of physics are expressed in terms of mathematics.

0.4 The oldest branch of mathematics is geometry; its earliest scientific use was in astronomy.

0.5 Geometry was formulated axiomatically by Euclid in his **Elements**.

0.6 Newton used that as a model in his **Principia** to formulate the laws of mechanics. Newton proved most of his results using geometric methods.

0.7 Einstein showed that gravity can be understood as the curvature of the geometry of space-time.

0.8 The other forces of nature (the electromagnetic, weak and strong forces) are also explained in terms of the geometry of connections on fibre bundles.

0.9 We dont yet have a unified theory of all forces; all attempts to construct such a theory are based on geometric ideas.

0.10 No one ignorant of geometry can be a physicist.

Chapter 1

Vector Spaces

1.1 A *vector space* is a set on which the operations of addition and multiplication by a number are defined.

1.2 The numbers can be *real* or *complex*; then we get *real* or *complex vector spaces* respectively. We will (unless said otherwise) work with real vector spaces.

1.3 The elements of a vector space are called *vectors*; the numbers we multiply vectors with are often called *scalars*.

1.4 Addition of vectors must be commutative and associative; there must be a vector 0 which when added to any vector produces itself.

1.5 Multiplication of a vector by a scalar must be distributive with respect to vector addition as well as scalar addition.

1.6 The smallest vector space is the set consisting of just the zero vector; this is called the trivial vector space and is denoted by 0.

1.7 The set of real numbers is itself a real vector space, called R.

1.8 The set of ordered n-tuples of real numbers is a vector space, with addition being defined component-wise and the scalar multiplying each component.

1.9 The above vector space is said to have dimension n; we will see an abstract definition of dimension later.

1.10 There are also vector spaces of infinite dimension; the set of all real valued functions on any set is a vector space. The dimension of this vector space is the cardinality of the set.

1.11 A map $f : V \to W$ between vector spaces is linear if $f(\alpha u + \beta v) = \alpha f(u) + \beta f(v)$.

1.12 If a linear map is one-one and onto, it is an *isomorphism*; the corresponding vector spaces are said to be *isomorphic*, $V \sim W$. 'Isomorphic' means 'having the same structure'.

1.13 The set V' of linear maps of a vector space V to R is called its dual. V' is also, of course, a vector space. The elements of V' are often called 1 -forms.

1.14 A *linear operator* is a linear map from V to itself. It makes sense to multiply linear operators: $LM(u) = L(M(u))$. Operator multiplication is associative but not always commutative.

1.15 An *algebra* is a vector space along with a bilinear multiplication $V \times V \to V$;i.e., $(\alpha u + \beta v)w = \alpha uw + \beta vw, u(\alpha v + \beta w) = \alpha uv + \beta uw.$

1.16 An algebra is *commutative* if $uv = vu$ for all pairs; it is *associative* if $u(vw) = (uv)w$ for all triples.

1.17 The set of linear operators on a vector space is an associative algebra; it is commutative only when the the vector space is either 0 or R .

1.18 There are two ways of combining two vector spaces to get a new one: the direct sum and the direct product.

1.19 The *direct sum* $V \oplus W$ of two vector spaces V and W is the set of ordered pairs (v, w) with the obvious addition and scalar multiplication : $(v, w) + (v', w') = (v + v', w + w')$ and $\alpha(v, w) = (\alpha v, \alpha w)$.

1.20 Transposition, $(v, w) \mapsto (w, v)$, is a natural isomorphism between $V \oplus W$ and $W \oplus V$; also $U \oplus (V \oplus W)$ and $(U \oplus V) \oplus W$ can both be thought of as the space of triples (u, v, w) . Hence we will just write $U \oplus V \oplus W$ etc.

1.21 It is clear that $R^{m+n} = R^m \oplus R^n$.

1.22 The *direct* or *tensor* product $V \otimes W$ of two vector spaces is the set of linear maps from V' to W .

1.23 Again $V \otimes W \sim W \otimes V$; $U \otimes (V \otimes W) \sim (U \otimes V) \otimes W$.

1.24 Also, $V \times R = V$; $R^m \otimes R^n = R^{mn}$.

1.25 We can iterate the direct product n times to get $V^{\otimes n} = V \otimes V \otimes \cdots V$. Its elements are called *contravariant tensors of order* n .

1.26 Since $V^{\otimes m} \otimes V^{\otimes n} = V^{\otimes (m+n)}$ it is natural to define $V^{\otimes 0} = R; V^{\otimes 1} = V$. Thus scalars are tensors of order zero while vectors are contravariant tensors of order one.

1.27 We can then take the direct sum of all of these to get the *total tensor space* $\mathcal{T}(V) = \otimes_{n=0}^{\infty} V^{\otimes n} = R \oplus V \oplus V \otimes V \cdots$.

1.28 $V'^{\otimes n}$ can be viewed also as the space of multilinear functions of n vectors. Its elements are also called *covariant tensors*.

1.29 The *direct product* of two 1-forms $u' \otimes v'$ can defined by its action on a pair of vectors: $u' \otimes v'(u, v) = u'(u)v'(v)$. A general element of $V' \otimes V'$ is a linear combination of such 'factorizable' elements.

1.30 More generally we can define the direct product of two covariant tensors of order m and n :

$$t \otimes \tilde{t}(u_1, \cdots u_m, v_1, \cdots v_n) = t(u_1, \cdots, u_m)\tilde{t}(v_1, \cdots v_n).$$

This turns $\mathcal{T}(V')$ into an associative but in general not commutative algebra. It is commutative only if $V \sim 0, R$.

1.31 An element $t \in V'^{\otimes n}$ is symmetric if it is invariant under permutations; e.g., $t(u, v) = t(v, u)$. The subspace of symmetric tensors is denoted by $\mathcal{S}^n(V')$ and $\mathcal{S}(V') = \bigoplus_{n=0}^{\infty} \mathcal{S}^n(V')$.

1.32 Averaging over all possible orderings gives a projection $\sigma : V'^{\otimes n} \to \mathcal{S}^n(V')$; e.g., $\sigma(t)(u, v) = \frac{1}{2}[t(u, v) + t(v, u)]$.

1.33 We can define a 'symmetrized multiplication' $s\tilde{s}$ of two symmetric tensors: $s\tilde{s} = \sigma(s \otimes \tilde{s})$. This turns $\mathcal{S}(V')$ into a commutative associative algebra.

1.34 A tensor $t \in V'^{\otimes n}$ is *antisymmetric* if it changes sign under an odd permutation of its arguments; e.g., $t(u,v) = -t(v,u)$. A covariant anti-symmetric tensor is also called a *form*; the space of anti-symmetric tensors is denoted by $\Lambda^n(V')$.

1.35 Averaging all permutations weighted by the sign of the permutation gives a projection $\lambda : V'^{\otimes n} \to \Lambda^n(V')$ to the space of anti-symmetric tensors.

1.36 The *wedge product* or *exterior product* of forms is defined by $a \wedge b = \lambda(a \otimes b)$. It is skew-commutative: $a \wedge b = (-1)^{mn} b \wedge a$ if $a \in \Lambda^m(V')$ and $b \in \Lambda^n(V')$.

1.37 The wedge product turns $\Lambda(V') = \oplus_n \Lambda^n(V')$ into an associative algebra called the *exterior algebra*.

1.38 An *inner product* $< .,. >$ on V is a symmetric bilinear map $V \times V \to R$ which is non-degenerate; i.e., $< u,v >=< v,u >$ and $< u,v >= 0 \forall u \Rightarrow v = 0$. It is positive if $< u,u > \geq 0$.

1.39 Clearly an inner product is a kind of covariant tensor; it is called the metric tensor. We may denote $< u,v >= g(u,v)$.

1.40 A positive inner product gives a notion of length for every vector; the square of the *length* of a vector is just its inner product with itself.

1.41 An inner product can be thought of as an invertible linear map from V to V'; i.e., it is an isomorphism between V and V'.

1.42 The inverse of the above map gives an inner product on V' given one on V.

1.43 The simplest example of an inner product is $< u,v >= \sum_i u^i v^i$ in R^n. The length of a vector is just its Euclidean distance from the center.

1.44 A *symplectic form* is an anti-symmetric non-degenerate bilinear map $V \times V \to R$. A vector space along with a symplectic form is a symplectic vector space.

1.45 Just as the metric tensor measures the length of a vector, the symplectic form measures the area of the parallelogram formed by a pair of vectors. The sign of the area contains information about the relative orientation of the vectors: if we reverse the direction of one of them the area changes sign.

1.46 A *complex structure* is a linear map $J : V \to V$ which satisfies $J^2 = -1$, where 1 denotes the identity map.

1.47 Let $\zeta = \alpha + i\beta$ be a complex number. Define $\zeta u = \alpha u + \beta Jv$. This turns a real vector space with a complex structure into a complex vector space. Every complex vector space can be obtained this way.

1.48 The definition of inner product on complex vector spaces is a bit different. We require the complex number $< u, v >$ to be *hermitean* rather than symmetric: $< u, v >=< v, u >^*$. Moreover it is linear in the second argument but *anti-linear* in the first : $< \alpha u + \beta v, w >= \alpha^* < u, w > +\beta^* < v, w >$.

1.49 A complex vector space with an inner product can be thought of as a real vector space with a complex structure and an inner product, with the inner product satisfying the additional condition that $\omega(u, v) = g(u, Jv)$ is antisymmetric; i.e., ω is a symplectic form.

1.50 Conversely given a complex structure J and a symplectic form ω we have a hermitean inner product if $g(u, v) = \omega(Ju, v)$ is symmetric.

1.51 The elements of the space $V^{\otimes m} \otimes V'^{\otimes n}$ are called *tensors of order* (m, n) . For example, a complex structure is a tensor of order $(1, 1)$.

1.52 The multiplication rule of an algebra can also be viewed as a tensor . For, a bilinear map $V \times V \to V$ can be viewed as a trilinear map $m : V' \times V \times V \to R$. The *structure tensor* m is an element of $V \otimes V'^{\otimes 2}$; i.e., a tensor of order $(1, 2)$.

Chapter 2

Index Notation for Tensors

2.1 Just as it is useful to represent numbers in the decimal (or other base) to perform calculations, it is useful to represent vectors by their components with respect to a basis.

2.2 A set of vectors $e_1, \cdots e_k \in V$ is *linearly independent* if $\sum_i \alpha_i e_i = 0 \Rightarrow \alpha_i = 0$.

2.3 A *basis* is a maximal set of linearly independent vectors.

2.4 That is, any vector $v \in V$ can be expressed as $v = \sum_i v^i e_i$ for some unique n -tuple of real numbers $(v^1, \cdots v^n)$. These numbers are called the *components of* v *with respect to* e_i .

2.5 We will see soon why we place the indices on the components as superscripts and not subscripts.

2.6 The maximum number of linearly independent vectors in V is called the *dimension* of V . We will consider the case of finite dimensional vector spaces unless stated otherwise.

2.7 Two bases $e_i, \tilde{e}_i$ are related by a linear transformation $\tilde{e}_i = A_i^j e_j$. A_i^j are the components of an invertible matrix.

2.8 The components of a vector with respect to these bases are also related by a linear transformation: $v = \sum_i v^i e_i = \sum_i \tilde{v}^i \tilde{e}_i, v^i = \sum_j A_j^i \tilde{v}^j$.

2.9 To simplify notation we will often drop the summation symbol. Any index that appears more than once in a factor will be assumed to be summed over. This convention was introduced by Einstein.

2.10 It will be convenient to introduce the Kronecker delta symbol: $\delta^i_j = 0$ if $i \neq j$ and $\delta^i_j = 0$ if $i = j$.

2.11 Given a basis e_i in V , we can construct its *dual basis* in V' : $e'^i(e_j) = \delta^i_j$.

2.12 A form can be expanded in terms of this dual basis: $\phi = \phi_i e'^i$. These components transform in the opposite way to the components of a vector: $\tilde{\phi}_i = A^j_i \phi_j$.

2.13 Although the components depend on the basis, the scalar $\phi(u) = \phi_i u^i$ is invariant under changes of basis.

2.14 More generally, the collection $e_{i_1} \otimes \cdots e_{i_m} \otimes e'^{j_1} \otimes \cdots e'^{j_n}$ is a basis in the space of tensors of order (m, n) : $t = t^{i_1 \cdots i_m}_{j_1 \cdots j_n} e_{i_1} \otimes \cdots e_{i_m} \otimes e'^{j_1} \otimes \cdots e'^{j_n}$.

2.15 Each upper index transforms contravariantly under changes of basis; each lower index transforms covariantly. This is the reason for distinguishing between the two types of indices. A contraction of an upper index with a lower index is invariant under changes of basis.

2.16 The metric tensor of an inner product space has components that satisfy $g_{ij} = g_{ji}$. Moreover, the determinant of the matrix g_{ij} is non-zero.

2.17 Given an inner product in V , we can construct one on V' ; its components g^{ij} form the matrix inverse to g_{ij} : $g^{ij} g_{jk} = \delta^i_k$.

2.18 We call the g^{ij} the contravariant components of the metric tensor and g_{ij} the covariant components.

2.19 Because the metric tensor can be thought of as an isomorphism between a vector space and its dual, it can be used to convert a contravariant index into a covariant index and vice-versa: $v^i \to g_{ij} v^j$ or $w_i \to g^{ij} w_j$. This is called lowering and raising of indices.

2.20 A symplectic structure ω has components ω_{ij} which form an invertible anti-symmetric matrix. Since the determinant of an odd dimensional anti-symmetric matrix is zero, symplectic forms exist only in even dimensional vector spaces.

2.21 A complex structure J has components J^i_j satisfying the condition $J^i_j J^j_k = -\delta^i_k$.

2.22 The multiplication rule of an algebra can be thought of as a tensor of order $(1,2)$. It has components m^i_{jk}; the components of the product of u and v is $m^i_{jk}u^j v^k$. The m^i_{jk} are called the *structure constants* of the algebra.

2.23 If the algebra is commutative, the structure constants are symmetric in the two lower indices: $m^i_{jk} = m^i_{kj}$. If it is associative, it satisfies the quadratic relation $m^p_{ij}m^q_{pk} = m^q_{ip}m^p_{jk}$.

2.24 A symmetric tensor $t \in \mathcal{S}^n(V')$ can be thought of as defining a polynomial of order n in the components of a vector: $t(v, v, \cdots v) = t_{i_1 i_2 \cdots i_n} v^i_1 \cdots v^{i_n}$. In fact this polynomial completely determines the symemtric tensor. The mutplication law for symmetric tensors we defined earlier is just the ordinary multiplication of polynomials: this explains why it is commutative.

2.25 The wedge product of forms can be thought of in a similar way as the polynomials in some 'anti-commuting' variable. This is a suggestive notation, particularly favored by physicists.

2.26 We define some abstract variables (*Grassmann variables*) ψ^i satisfying $\psi^i\psi^j = -\psi^j\psi^i$ for all i, j. In particular $(\psi^i)^2 = 0$. A polynomial or order n in these variables will be $t_{i_1 \cdots i_n}\psi^{i_1} \cdots \psi^{i_n}$. The multiplication of these polynomials is equivalent to the wedge prodcut of forms.

2.27 More precisely, the exterior algebra is the free algebra generated by the ψ^i quotiented by the ideal generated by the relations $\psi^i\psi^j = -\psi^j\psi^i$.

2.28 The free algebra on n variables is just the tensor algebra on an n-dimensional vector space.

2.29 No physicist can work with symmetric and anti-symmetric tensors without being reminded of bosonic and fermionic quantum systems. Indeed if V is the space of states of a particle, the space of states of an n particle system is $\mathcal{S}^n(V)$ for bosons and $\Lambda^n(V)$ for fermions. The only point to remember is that the vector spaces of interest in quantum mechanics are over the field of complex numbers.

2.30 The components of a vector can be thought of as the Cartesian coordinates on V . Given a function $f : V \to R$ and a basis e_i in V we can construct a function $\tilde{f} : R^n \to R, \tilde{f}(x_1, x_2, \cdots x_n) = f(x^i e_i)$. We can now define a f to be *differentiable* if $\tilde{f}$ has continuous partial derivatives as a function on R^n .This class of functions is called $C^1(V)$.

2.31 Analogously we define $C^k(V)$, $C^\infty(V)$, $C^\omega(V)$. They are respectively, the class of functions whose k th derivatives are continuous, then those with continuous partial derivatives of any order and finally functions which have a convergent Taylor series expansion. C^∞ functions are called *smooth* and the C^ω are *analytic*.

2.32 Analytic functions are completely defined by their derivatives at the origin; these derivatives are symmetric tensors. Thus, each element of $\mathcal{T}(V)$ for which the infinite series $\sum_{n=0}^{\infty} t_{i_1 \cdots i_n} x^{i_1} \cdots x^{i_n}$ converges characterizes an analytic function. There is no similar characterization of smooth or differentiable functions. For polynomials this series terminates.

Chapter 3

Curvilinear Co-ordinates

3.1 A basis in a vector space is also a co-ordinate system: it associates a set on $n = \dim V$ numbers with every point that uniquely specifies it.

3.2 Often it is more convenient to use co-ordinate systems that are not linear even though the underlying space is linear; e.g., when the system has spherical symmetry we use spherical polar co-ordinate system.

3.3 When the underlying space itself is curved (e.g., a sphere) we have no choice but to use curvilinear co-ordinates.

3.4 Polar co-ordinates on the plane are the simplest example: Each point $(x^1, x^2) \in R^2$ is specified by its distance r from the origin and the angle θ that the radial vector makes with the first axis. In other words, $x^1 = r\cos\theta, x^2 = r\sin\theta$.

3.5 It is useful to have a formula for the metric in curvilinear co-ordinates. An infinitesimal displacement (dx^1, dx^2) has length^2 equal to $ds^2 = [dx^1]^2 + [dx^2]^2$. Transforming with the above formula we get $ds^2 = dr^2 + r^2 d\theta^2$.

3.6 In R^3 , we have the spherical polar co-ordinates in which $ds^2 = dr^2 + r^2[d\theta^2 + \sin^2\theta d\phi^2$].

3.7 A *co-ordinate system* on an open neighborhood U of R^n is a smooth injective map $x : U \to R^n$. Injective means that each point is uniquesly specified by the value of the co-ordinates at that point.

3.8 If $x^1, \cdots x^n$ are the usual Cartesian co-ordinates, a general co-ordinate system y is given by smooth functions $y^i = f^i(x)$ which have non-vanishing Jacobian $\det \frac{\partial y}{\partial x}$ in U .

3.9 The metric $ds^2 = \delta_{ij} dx^i dx^j$ becomes $ds^2 = g_{ij}(y) dy^i dy^j$

where $g_{ij}(y) = \frac{\partial y^k}{\partial x^i} \frac{\partial y^l}{\partial x^j} \delta_{kl}$.

3.10 The Jacobian of the transformation is the square root of the determinant of g_{ij} ; i.e., $\det \frac{\partial y}{\partial x} = \sqrt{[\det g]}$.

3.11 The gradient of a function transforms covariantly under changes of co-ordinates: $\frac{\partial V}{\partial y^i} = \frac{\partial V}{\partial x^j} \frac{\partial x^j}{\partial y^i}$.

3.12 A *curve* is on R^n is any smooth function $f : R \to R^n$. At each point on the curve, the derivative $\frac{df(t)}{dt}$ is its *tangent vector*.

3.13 If we make a change of co-ordinates the components of a tangent vector transform contravariantly.

3.14 It is useful to have formulae for some standard operators such as the Laplace operator in curvilinear co-ordinates. It is actually simplest to work out the general case first.

3.15 The Poisson equation $\nabla^2 V = \rho$ is the condition for the 'electrostatic energy' $\frac{1}{2} \int [\nabla V]^2 d^3x + \int \rho(x) V(x) d^3x$ to be a minimum.

3.16 We can get a formula for the Laplace operator by first working out a formula for the energy.

3.17 The energy is $\frac{1}{2} \int \sqrt{g} g^{ij} \frac{\partial V}{\partial y^i} \frac{\partial V}{\partial y^j} d^3y + \int \rho V \sqrt{g} d^3y$. By varying this we get the Poisson equation $\frac{1}{\sqrt{g}} \partial_i \left[\sqrt{g} g^{ij} \partial_j V \right] = \rho$. We are using g as an abbreviation for $\det g$.

3.18 Thus we get the formula for the Laplacian in an arbitrary co-ordinate system: $\Delta V = \frac{1}{\sqrt{g}} \partial_i \left[\sqrt{g} g^{ij} \partial_j V \right]$. We will see that this formula is valid even for curved spaces.

3.19 There are some co-ordinate systems in which the Laplace equation is separable. This allows its analytic solution when the boundary conditions have a simple form in one of these co-ordinates.

3.20 The solution in cartesian co-ordinates, spherical polar co-ordinates, cylindrical polar co-ordinates etc. are well-known special cases. You should work out the formula for the Laplacian in these co-ordinates.

3.21 In fact they are all special cases of a remarkable co-ordinate system of Jacobi-the confocal ellipsoidal co-ordinates. See *Higher Transcendental Functions* Vol III, ed. by Erdelyi. We will digress to describe these co-ordinates; this part is meant for the more advanced students.

3.22 Let $a > b > c$ be three real numbers; let (x, y, z) be the usual Cartesian co-ordinate system in R^3 . Consider the equation

$$\frac{x^2}{a+\theta} + \frac{y^2}{b+\theta} + \frac{z^2}{c+\theta} = 1.$$

For each point in the quadrant R^3 such that $x, y, z > 0$ there are three solutions to this equation for θ ; it can be rewritten as a cubic equation for θ with discrimant proportional to xyz . There will be one solution in each of the intervals $(-c^2, \infty), (-b^2, -c^2), (-a^2, --b^2)$. Jacobi's idea is to use these three solutions (λ, μ, ν) as the co-ordinates of the point.

3.23 The surface λ =constant is an ellipsoid, that with μ =constant is a hyperboloid of one sheet, and ν =constant gives a hyperboloid of two sheets. They all have the same pair of points as the foci: this is a confocal family of quadrics.

3.24 The remarkable fact is that the Laplace equation is separable in these co-ordinates. Thus they are appropriate to solve electrostatic problems in which boundary conditions are given on a quadric. The solutions involve the Lamé functions.

3.25 Some remarkable results can be proved this way. For example, the electrostatic field outside an ellipsoid of constant charge density is the same as that of a point charge at its center; the field inside is zero. This is a theorem of Newton and Ivory which can be proved also by an elementary, but clever, argument. See S. Chandrashekhar, *Ellipsoidal Figures of Equilibrium*

3.26 The Jacobi co-ordinates generalize to Euclidean space of any dimension: given numbers $a_1 > a_2 > \cdots > a_n$, we have the confocal quadrics

$$\sum_i \frac{x_i^2}{a_i^2 + \lambda} = 1.$$

There are n solutions λ_i which can again be used as co-ordinates. The Laplace equation is separable in these co-ordinates in any dimensions .

3.27 The geodesic equation on the surface of an ellipsoid of any dimension can be exactly solved by using these ideas: the Hamilton-Jacobi equation is separable. The solution involves the Jacobi elliptic functions.

Chapter 4

The Sphere and the Hyperboloid

4.1 We now begin the study of manifolds. We will look at a few examples before presenting a definition of a manifold.

4.2 Next to the plane, the most familiar surface is a sphere. The surface of the earth approximates a sphere. The sky can also be thought of as a sphere, the celestial sphere.

4.3 The geometry of spheres was first developed by ancient astronomers. The **Aryabhatiya** (written in 499 A.D.) contains results on non-Euclidean geometry.

4.4 The two dimensional *sphere* S^2 is the set of points at unit distance from the origin in R^3 .

4.5 From each point in S^2 we can draw a line connecting it to the origin. The distance between two points can be defined to be the angle (in radians) at which these radial lines intersect at the origin.

4.6 If two points are at zero distance they are identical. The maximum distance between points is π ; this happens when the points are *anti-podal* to each other; i.e., when the straight-line connecting them passes through the center of the sphere.

4.7 A *great circle* on S^2 is the intersection of the sphere with a plane that passes through the center. Any two distinct points lie along a great circle; for these two points and the center define a plane in R^3 (if the points are not antipodal). If the points are antipodal they lie on an infinite number of great circles.

4.8 If two points are not antipodal, (the short arc of) the great circle is the shortest path on the sphere between the two points. Such curves are called *geodesics*.

4.9 Thus longitudes (also called meridians) are geodesics but the only latitude (or parallel) that is a geodesic is the equator.

4.10 A *spherical triangle* consists of three distinct points connected pairwise by geodesics. The angle at each vertex is defined to be the angle between the tangents at that point. The sum of these angles need not be π ; indeed it is always greater than π . This shows that the axioms of Euclid cannot hold for spherical geometry.

4.11 For example we can draw a spherical triangle connecting the north pole to two points on the equator where each of the angles is a right angle.

4.12 In the *Aryabhatiya* the following problem is solved: given two sides of a spherical triangle and the angle at which they intersect, determine the remaining side and angles. Can you solve this?.

4.13 Geodesics are the closest analogues to straight lines on the surface of a sphere. But there are important differences between spherical and Euclidean (planar) geometry: any two geodesics will meet.

4.14 An axiom of Euclidean geometry is that *given a straight line L and a point P not on it, there is exactly one straight line passing through P but not intersecting L .* This axiom does not hold in spherical geometry. This is due to the curvature of the surface of the sphere.

4.15 Geodesics on a sphere tend to 'focus': even though two geodesics starting at a point will move away from each other at first, they will eventually converge and meet at the anti-podal point. This is the feature of positive curvature.

4.16 A plane that intersects the sphere at exactly one point is called the *tangent plane* at that point. Being a vector space it is also called the *tangent space* . The tangent plane is the best planar approximation of the sphere at that point.

4.17 The tangent space at $p \in S^2$ is also the set of all vectors in R^3 orthogonal to the radius vector at p .

4.18 Elements of the tangent space at p are called *tangent vectors*.

4.19 The *length* of a tangent vector is just its length thought as an element of R^3 . Thus on each tangent space of S^2 there is a metric tensor, called the *induced metric* .

4.20 A curve in S^2 is a smooth function $f : R \to R^3$ that lies on the sphere: $< f(t), f(t) >= 1$.

4.21 The derivative of $f(t)$ is a vector in R^3 orthogonal to $f(t)$; hence it can be thought of as an element of the tangent space at $f(t)$. This is called the *tangent vector* to the curve.

4.22 The *length* of a curve is the integral of the lengths of its tangent vectors: $l(f) = \int_R \sqrt{< \dot{f}(t), \dot{f}(t) >}\, dt$. (The dot denotes differentiation with respect to t .)

4.23 The shortest curve connecting two points is a geodesic; the length of the geodesic is the distance between points as defined previously. These facts can be proved using variational calculus.

4.24 A function $f : S^2 \to R$ is *smooth* if it is the restriction to the sphere of some smooth function on R^3 . The set of smooth functions is a *commutative ring*: it is closed under addition and multiplication. Division is not well-defined since functions may vanish.

4.25 The set of all tangent vectors at the various points on the sphere is called its *tangent bundle* $TS^2 = \{(r, u) \in R^3 \times R^3 | r^2 = 1; r \cdot u = 0\}$.

4.26 A *vector field* on the sphere is a smooth function that assigns to each point in S^2 a tangent vector at that point. More precisely, it is a smooth function $u : S^2 \to R^3$ such that $r \cdot u(r) = 0$.

4.27 Given any non-zero constant $a \in R^3$, its cross product with the radius vector gives a vector field on the sphere. Even if $a \neq 0$, the vector field $a \times r$ vanishes at two points on the sphere: $\frac{a}{|a|}$ and its antipodal point, $-\frac{a}{|a|}$.

4.28 In fact this is a general fact: any smooth vector field on the sphere vanishes somewhere. We will not prove this statement in this course.

4.29 A vector field can be thought of as an infinitesimal displacement of each point. It is useful to picture it as the velocity of some fluid attached to the surface of the sphere.

4.30 A vector field can also be thought of as a first order differential operator acting on the ring of smooth functions. We will examine this point of view later.

4.31 It is useful to establish a co-ordinate system: a pair of real numbers that will specify a point on the sphere.

4.32 An example is the spherical polar co-ordinate system. We fix a point ('North Pole') and a geodesic ('standard meridian') that connects it to its antipode ('South Pole'). The *latitude* of point θ is its distance from the origin; i.e, the length of the geodesic connecting the point to the North Pole. The *longitude* ϕ is the angle this geodesic makes with the standard meridian at the North Pole.

4.33 Excepting the points on the standard meridian, each point is uniquely specified by the ordered pair (θ, ϕ) . There is no single co-ordinate system that can cover all of the sphere.

4.34 A curve (that doesn't intersect the meridian) is given by two functions $\theta(t)$ and $\phi(t)$. Its length is given by $\int \sqrt{[\dot{\theta}^2(t) + \sin^2\theta(t)\dot{\phi}^2(t)]}dt$.

4.35 The *hyperboloid* is the set of points in R^3 satisfying $z^2 - x^2 - y^2 = 1; z > 0$.

4.36 It is possible to define the notions of tangent space, vector field, curves, length of curves and geodesics for this case by analogy to the sphere.

4.37 The resulting geometry (studied first by Lobachewsky, Bolyai and Poincare) has negative curvature; two geodesics emanating from a point diverge from each other and never meet.

4.38 We will return to these examples after developing the general theory of manifolds.

Chapter 5

Vector Fields

5.1 A *curve* is a smooth function $\gamma : [a, b] \to M$.

5.2 Given any co-ordinate system in a neighborhood $U \subset M$, $x : U \to R^n$, we can think of a curve as a function $x \circ \gamma : [a, b] \to R^n$.

5.3 Consider two curves that pass through the same point $\gamma(t_1) = \tilde{\gamma}(t_2) = p$. We say that they are *equivalent to first order* at p if in any coördinate system, $\lim_{t\to t_1} \frac{x\circ\gamma(t)-x(p)}{t-t_1} = \lim_{t\to t_2} \frac{x\circ\tilde{\gamma}(t)-x(p)}{t-t_2}$ .This relation is independent of the coördinate system.

5.4 A class of curves equivalent to first order at p is a *tangent vector* at p . There is a natural addition and multiplication by scalars on the set of tangent vectors which makes it a real vector space of dimension n , the tangent space T_pM to M at p .

5.5 The point is that a curve is determined to first order in t by its tangent vector. So we turn this idea around and define a tangent vector as the equivalence class of curves that agree to first order.

5.6 Given a coördinate system x , there is a basis for T_pM . The basis vectors are the tangents to the coördinate axes: the curves we get by varying just one coördinate at a time.The tangent vector to the curve γ has components $\lim_{t\to t_1} \frac{x\circ\gamma(t)-x(p)}{t-t_1}$.

5.7 A *derivation* of a ring is a map V such that

$$V(f + g) = Vf + Vg, \quad V(fg) = (Vf)g + fVg.$$

5.8 A derivation can be thought of as an infinitesimal isomorphism of a ring. More precisely, let ϕ_t^* be a one-parameter family of isomorphisms of a ring such that ϕ_0 is the identity. Then the equation $\phi_t^*(fg) = (\phi_t^* f)(\phi_t^* g)$ implies $\frac{d}{dt}\phi_t^*(fg)|_{t=0} = (\frac{d}{dt}\phi_t^* f)|_{t=0} g + f\frac{d}{dt}\phi_t^* g|_{t=0}$.

5.9 For the ring $C(R)$ of smooth functions on the real line, the derivations are the first order differential operators $Vf(x) = v(x)\frac{df}{dx}$. For $C(R^n)$, also, any derivation is of the form $Vf(x) = v^i(x)\frac{\partial f}{\partial x^i}$. Thus derivations of $C(R^n)$ correspond to vector fields on R^n , the components of which are $v^i(x)$.

5.10 A *vector field* on a manifold M is a derivation of its coördinate ring $C(M)$.

5.11 In a coördinate system, a vector field is a first order differential operator $X = X^i(x)\frac{\partial}{\partial x^i}$. The functions X^i are the *components* of X with respect to the coördinate basis $\frac{\partial}{\partial x^i}$.

5.12 A vector field can be thought of as an infinitesimal diffeomorphism of the manifold to itself. Suppose $\phi : [a, b] \times M \to M$, so that $(t, p) \mapsto \phi_t(p)$ is a diffeomorphism for each value of t , and that $\phi_0(p) = p$. This is called a one-parameter family of diffeomorphisms. The infinitesimal effect of this diffeomorphism on functions is the derivation $Xf(p) = \frac{d}{dt} f(\phi_t(p))|_{t=0}$.

5.13 If we hold p fixed and vary t , $\phi_t(p)$ is a curve passing through p ; at p , the vector field X can be identified with the tangent vector to this curve.

5.14 Conversely, given a vector field X , there is a one-parameter family of diffeomorphisms $\phi : [a, b] \times M \to M$ to which it is tangent, for some a and b . In terms of coördinates this means that the system

$$\frac{d\phi_t^i(x)}{dt} = X^i(\phi_t(x)) \quad \phi_{t=0}^i(x) = x^i$$

has a solution that is smooth, and that the functions ϕ_t have smooth inverses. This is a rather difficult theorem on the existence of solutions to ordinary differential equations.

5.15 The solutions to the above equation satisfy $\phi_t(\phi_s(x)) = \phi_{t+s}(x)$. Thus we can regard ϕ_t as depending exponentially on the parameter t . Indeed we can think of ϕ_t as $\exp tX$.

5.16 The curves $\phi_t(p)$ obtained by keeping p fixed and varying t are called the *integral curves* or *characteristic curves* of the vector field X. They can be immensely complicated even for innocent looking vector fields. A whole field of mathematical physics called *chaos* has developed around understanding the long time behavior of integral curves.

5.17 A simple example (not chaotic) is a rotation in a plane:

$$\phi_t(x, y) = (x \cos t - y \sin t, x \sin t + y \cos t).$$

This is the integral curve of the vector field $X = -y\frac{\partial}{\partial x} + x\frac{\partial}{\partial y}$.

5.18 Consider the torus $S^1 \times S^1$ and the curve $\gamma(t) = (t, \omega t)$, where r is some real number. If ω is a rational number $\omega = \frac{p}{q}$, this is a curve that winds around the torus and returns to the starting point in a time $2\pi q$: it will make q revolutions around the first circle and p revolutions around the second. If ω is irrational, this curve will never return to the starting point. But it will come arbitrarily close to *any* point on the torus. The integral curve fills the whole torus!. This shows how very simple vector fields can have complicated integral curves.

5.19 The book *Differential Equations: Geometric Theory* by S. Lefshetz is a classic text on the subject of integral curves. We will come back to this topic in the context of classical mechanics.

5.20 The product of two derivations is in general not a derivation; however the commutator of two derivations $[U, V]f = U(Vf) - V(Uf)$ is always a derivation. This allows us to define the commutator of vector fields.

5.21 In coördinates $[U, V]f = \left(u^k \frac{\partial v^i}{\partial x^k} - v^k \frac{\partial u^i}{\partial x^k}\right) \frac{\partial f}{\partial x^i}$. The point is that the product of two first order differential operators is in general of second order; but in the commutator the second order terms cancel out so we again get a first order operator.

5.22 The set of vector fields form a Lie algebra under this operation; i.e., $[U, V] = -[V, U]$ and $[U, [V, W]] + [V, [W, U]] + [W, [U, V]] = 0$.

5.23 An infinitesimal rotation around any axis is a vector field on S^2. If we denote the infinitesimal rotation around the first axis by L_1, that around the second by L_2 and so on, we can check that $[L_1, L_2] = L_3, [L_2, L_3] = L_1, [L_3, L_1] = L_2$.

5.24 A smooth function on S^1 can be expanded in a Fourier series $f(\theta) = \sum_n f_n e^{in\theta}$. The coefficients vanish faster than any power of n: $\lim_{n\to\infty} |n|^k |f_n| = 0$. Any vector field can also be expanded: $V = \sum_n v_n e^{in\theta} \frac{d}{d\theta}$. The vector fields $L_n = e^{in\theta} \frac{d}{d\theta}$ form a basis. They satisfy the commutation relation $[L_m, L_n] = i(n-m)L_{m+n}$.

5.25 The sum of vector fields is a vector field; so is its product with a real number.

5.26 We can multiply a vector field *on the left* by a smooth function to get another vector field. The components get multiplied by the function. Thus the set of vector fields is a module over the ring of functions.

5.27 Given a coördinate system x^i on an open neighborhood U, there is a basis of vector fields in U: $\partial_i = \frac{\partial}{\partial x^i}$. These vector fields commute with each other: $[\partial_i, \partial_j] = 0$. Conversely any set of linearly independent (over $C(M)$) vector fields that commute provide a co-ordinate system in some neighborhood. We will see this when we talk about integral curves.

5.28 A smooth map $\phi : [a, b] \times M \to M$ is a one-parameter family of diffeomeorphisms if ϕ_t is a diffeomorphism for each t and $\phi_t \circ \phi_u = \phi_{t+u}$.

Chapter 6

Differential Forms

6.1 Notice that the division operation on scalars is never used in the axioms defining a vector space. The same axioms where the axioms take values in a ring is called a *module*. Its elements are called *vector fields*.

6.2 The dual of a module, tensor products etc. are all defined analogously. There are some important differencers however: there may not exist a basis for example.

6.3 The space of vector fields on a manifold $V(M)$ is a module over the ring of smooth functions $C(M)$. Any vector field can be multiplied on the left by a function: $(fX)(g) = f(Xg)$ in a way that satisfies all the axioms of a module.

6.4 The dual of the module of vector fields is the space of *one-forms* or *covector fields*. Thus a one-form $\omega : V(M) \to C(M)$ is a map that associates to every vector field v a function $\omega(v)$ such that $\omega(fv) = f\omega(v)$.

6.5 Sometimes $\omega(v)$ is also denoted by $i_v\omega$ and may be called the *contraction* of the pair (ω, v).

6.6 In an an analogous way we can define a covariant tensor field T of order r as a map, linear over $C(M)$, of an ordered set of r vector fields, $T(v_1, .. \cdots v_r)$. It is very important that the property of linearity $T(f_1v_1, f_2v_2, \cdots f_rv_r) = f_1f_2 \cdots f_rT(v_1, \cdots v_r)$ hold for arbitrary smooth functions $f_1, \cdots f_r)$ and not just for constants.

6.7 Contravariant tensor fields are $C(M)$ -linear maps of an ordered set of r one-forms; mixed tensor fields of order (r, s) depend linearly on r one-forms and s vector fields. Symmetric and antisymmetric tensor fields are defined as before for tensors.

6.8 Covariant anti-symmetric tensor fields are also called *differential forms*.

6.9 An example of a one-form is the *exterior derivative* of a function, defined by $df(X) = Xf$. Thus the exterior derivative is a generalization of the gradient of a function familiar from vector calculus.

6.10 Recall that there is an associative multiplication on the space $\mathcal{T}(M) = \oplus_{r=0}^{\infty}\mathcal{T}^r(M)$ of all covariant tensors, the direct product:

$$\alpha \otimes \beta(u_1, \cdots u_{r+s}) = \alpha(u_1, \cdots u_r)\beta(u_{r+1} \cdots u_{r+s}).$$

Also there is a projection map $\lambda : \mathcal{T}^r(M) \to \Lambda^r(M)$ which picks out the anti-symmetric part of a covariant tensor:

$$\lambda(\alpha)(u_1, \cdots u_r) = \frac{1}{r!}\sum_{\pi \in S_r} \text{sgn } (\pi)\alpha(u_{\pi(1)} \cdots u_{\pi(r)}).$$

6.11 Conbining these gives us a multiplication on the space of differential forms called the *exterior product* or the *wedge product*:

$$\alpha \wedge \beta = \lambda(\alpha \otimes \beta).$$

The wedge product is graded commutative; i.e.,

$$\alpha \in \Lambda^r(M), \beta \in \Lambda^r(M) \Rightarrow \alpha \wedge \beta \in \Lambda^{r+s}(M) \text{ and } \quad \alpha \wedge \beta = (-1)^r \beta \wedge \alpha$$

Moreover it is associative.

6.12 Every differential form $\alpha \in \Lambda^r(M)$ can be written in the form $\alpha = f_{i_1 \cdots i_r} dg^{i_1} \wedge dg^{i_r}$ for some functions $f_{i_1, \cdots i_r}, g_i \in C(M)$. The number of functions g^i needed to do this can exceed the dimension of the manifold; also, the differential form may not have a unique set of 'components' $f_{i_1, \cdots i_r}$ even for a fixed set of g_i . Within a coördinate neighborhood $U \subset M$, there is always a special class of functions, the coördinates x^i . A differential form can be expressed *uniquely* in terms of these coördinates:

$$\alpha = \alpha_{i_1 \cdots i_r} dx^{i_1} \wedge \cdots dx^{i_r}, \quad \alpha_{i_1 \cdots i_r} \in C(U).$$

6.13 The basis dx^i for one-forms is dual to the basis ∂_j for vector fields in U : $dx^i(\partial_j) = \delta^i_j$.

6.14 Differential forms are special among tensor fields in that there is a natural notion of differentiation , the *exterior derivative*, $d : \Lambda^r(M) \to \Lambda^{r+1}(M)$. It satisfies the axioms

$$d[\alpha + \beta] = d\alpha + d\beta, \quad d[\alpha \wedge \beta] = [d\alpha] \wedge \beta + (-1)^r \alpha \wedge d\beta, \quad d(d\alpha) = 0, \forall \alpha.$$

Moreover, on zero forms it is the same as the derivative defined earlier.

6.15 The explicit formula in terms of components follows easily

$$d\alpha = d\alpha_{i_1 \cdots i_r} \wedge dx^{i_1} \wedge \cdots dx^{i_r}$$

or even

$$d\alpha = \partial_{i_1} \alpha_{i_2 \cdots i_{r+1}} dx^{i_1} \wedge \cdots dx^{i_{r+1}}.$$

6.16 In electrodynamics, the magnetic field is the curl of the vector potential. In our language the magnetic field is really a two-form B in R^3 and its potential a one-form A , with $B = dA$. If we add the derivative of a scalar field to A , the magnetic field remains unchanged: $d[A + d\phi] = dA + d(d\phi) = dA$. This symmetry is called gauge invariance and is of enormous importance.

6.17 In relativistic langauge the electrostatic potential A_0 and the magnetic potential A above combine to a one-form in four-dimensional Minkowski space, $A = A_0 dx^0 + A_i dx^i$; the electric and magnetic fields combine to give a two form in Minkowski space $F = F_{0i} dx^0 \wedge dx^i + F_{ij} dx^i \wedge dx^j$. Here, $E_1 = F_{01}$ etc. and $B_1 = F_{23}$ etc. Then $F = dA$ and we still have a gauge invariance under $A \mapsto A + d\phi$.

6.18 The subset of Maxwell's equations that do not involve sources become now $dF = 0$; we can understand this as a necessary condition for the existence of a potential A such that $F = dA$. We will see later how to understand the rest of Maxwell equations in terms of differential forms.

Chapter 7

Thermodynamics

7.1 Thermodynamics is the study of heat.

7.2 Examples of systems to which thermodynamics can be applied are: a gas in a box of variable volume as in a steam engine ; a liquid which can evaporate to a gas as in a refrigerator; a mixture of chemicals; a magnet; a star; a blackhole.

7.3 The set of all possible states of a thermodynamic system is a manifold M , called the *state space*. The *thermodynamic observables* are functions on this manifold. Examples of thermodynamic observables are: the pressure exerted by the gas on the walls of its container and its volume; the volume of the gas and the liquid and their pressures; the magnetic moment ; the pressure, volume and relative abundances of various chemicals; the size, pressure, and the abundances of different nuclei in a star; the electric charge, angular momentum and mass of the blackhole.

7.4 In all these cases the system has many more degrees of freedom than the dimension of the thermodynamic state space. But it is too difficult to measure all these variables. Thermodynamics is an approximate or average description of the system in terms of a few degrees of freedom which have a macroscopic meaning.

7.5 In the case of a blackhole, the laws of physics (causality) prevents us from making measurements of the interior of the blackhole; the only degrees of freedom of the blackhole that can be measured are its electric charge, angular momentum and mass. Hence a thermodynamic description is not just a matter of conveinence; it is forced upon us.

7.6 When a system changes its thermodynamic state (eg., when a gas expands pushing on a moving piston in a cylinder) it can perform mechanical work.

7.7 Consider two systems which are isolated from the rest of the world but are in contact with each other. Eventually they come into *equilibrium*; i.e, their thermodynamic variables will become independent of time and they will stop exchanging energy. It is found emprically that *there is a thermodynamic observable called* temperature *such that two systems are in equilibrium if and only if they are at the same temperature*. This is called the *zeroth law* of thermodynamics.

7.8 Note that this law does not set a natural scale for temperature. Indeed any monotonic function $t \mapsto \phi(t)$ will give an equally good notion of temperature.

7.9 For example, two containers of ideal gases are in equilibrium if $P_1V_1 = P_2V_2$.

7.10 *The work done by isolated system in changing the state of a system from a state* p_1 *to another* p_2 *is independent of the path taken.* This is the *first law* of thermodynamics.

7.11 What we mean is that the system is isolated from other thermodynamic systems. It may be in contact with and exchange energy with a mechanical system.

7.12 This implies that there is a function $U : M \to R$ called *internal energy* such that the work done in changing from p_1 to p_2 is $U(p_2) - U(p_1)$. This function is defined only upto the addition of a constant.

7.13 When two systems are brought into contact with each other the internal energy of the total system is just the sum of their internal energies. This is not true in mechanics, due to the possibility of interaction between the systems; however such interaction energies are negligible in the thermodynamic approximation.

7.14 When a body is in contact with another, the work W done in changing its state from p_1 to p_2 depends in general on the path taken. The difference $Q = U - W$, is called *heat*.

7.15 If the change to p is infinitesimal, it is described by a vector v at p ; the infinitesimal amount of heat due to this change depends linearly on v . Hence there is a one-form ω , such that the heat emitted during the infinitesimal change v is $\omega(v)$. We will call this the *heat form*.

7.16 In most older textbooks this heat form is called dQ . But this notation is confusing since there is in general no function $Q : M \to R$ such that $\omega = dQ$: *the heat form is in general not an exact differential*.

7.17 Let us digress to establish a *theorem of Caratheodory* . This theorem will be useful in understanding the second law of thermodynamics. It is of intrinsic interest in geometry.

7.18 Given a one-form ω , let us say that a point $q \in M$ is *accesible* from $p \in M$ if there is a curve γ connecting them such that $\omega(\dot{\gamma}) = 0$. Otherwise q is *inaccessible* from p .

7.19 *Given a one-form ω , suppose that every neighborhood U of a point has inaccessible points. Then $\omega \wedge d\omega = 0$; morever, there are functions $T, S : U \to R$ such that $\omega = TdS$ in U .*

7.20 The function T is called an *integrating factor*. T and S are unique upto two constants: $T \mapsto k_1 T, S \mapsto k_1^{-1} S + k_2$, with $k_1 \neq 0$ gives the same ω .

7.21 Proof of Caratheodory's Theorem: The points accessible to p will form either an open set $V \subset U$ or a surface of some higher codimension (inverse function theorem of calculus). At each point p , there is a subspace of codimension one of the tangent space $T_p M$ of vectors transversal to ω : $\omega(X) = 0$. So the accessible points cannot form a surface of codimension greater than one. If there are inaccessible points in every neighborhood they cannot form an open subset either. Thus the points accesible from p must form a surface of codimension one in U . We can divide U into surfaces of points which are mutually accessible. Because they are of codimension one, there is a function S which is constant on these surfaces; i.e., $dS(X) = 0$ whenever $\omega(X) = 0$. Thus there is a never–vanishing function $T : U \to R$ such that $\omega = TdS$. Then $\omega \wedge d\omega = 0$.

7.22 Let us return to thermodunamics. An *adiabatic process* is a curve $\gamma : [a, b] \to M$ such that the heat form is zero on its tangent vectors: $\omega(\dot{\gamma}(t)) = 0$. Thus infinitesimal adiabatic processes are given by vectors X for which $\omega(X) = 0$. Thus two points are connected by an adiabatic process if they are accessible with respect to the heat form.

7.23 The second law of thermodynamics says that *in every neighborhood of a given state, there exist states that cannot be reached by an adiabatic process.*

7.24 From Caratheodory's theorem, this means that (at least locally) the heat form has an integrating factor: $\omega = TdS$. T is called *temperature* and S *entropy*. To justify the name temperature for T , we must prove that two systems are in equilibrium when they have the same value for T .

7.25 The temperature T does not have to be positive; however it cannot vanish so its sign cannot change. There are systems of negative temperature in nature: population inverted atomic systems for example.

7.26 Standard arguments in most thermodynamics books show that this form of the second law is equivalent to the other forms due for example to Kelvin or Carnot. We will not discuss them here.

Chapter 8

Classical Mechanics

8.1 Mechanics is the oldest and most fundamental branch of physics; the rest of physics is modelled on it. A natural formulation of the laws of mechanics is in terms of sympletic geometry.

8.2 A *symplectic form* ω on a manifold is a closed non-degenerate 2-form; i.e., such that $d\omega = 0$ and $\omega(., u) = 0 \Rightarrow u = 0$.

8.3 For such a form to exist, M must be of even dimension.

8.4 The basic example is $M = R^n \times R^n = \{(p, q)\}$, with $\omega = dp_i \wedge dq^i$. The coördinates p_i and q^i are said to be canonically conjugate to each other.

8.5 Given a function $f : M \to R$, we can define a vector field X_f by the relation $\omega(., X_f) = df$. It is called the *canonical vector field* of f .

8.6 The *Poisson bracket* of a pair of functions is $\{f, g\} = X_f g$. It satisfies

$$\{f, g\} = -\{g, f\}, \quad \{\{f, g\}, h\} + \{\{g, h\}, f\} + \{\{h, f\}, g\} = 0$$

and of course,

$$\{f, gh\} = \{f, g\}h + f\{g, h\}.$$

In other words it turns the space of functions into a Lie algebra; moreover, the Poisson bracket acts as a derivation on the multiplication of functions.

8.7 The set of all possible instantaneous states of a classical mechanical system is a symplectic manifold, called its *phase space*. An *observable* is a real valued function on the phase space. From an observable f we can construct its canonical vector field X_f .

8.8 Time evolution is given by the integral curves of the canonical vector field of a particular function, called the *hamiltonian*. Thus a classical mechanical system is defined by a triple (M, ω, H) : a differential manifold M , a symplectic form ω on it and a smooth function $H : M \to R$.

8.9 The total time derivative of any function is given by its Poisson bracket with the hamiltonian $\frac{df}{dt} = \{H, f\}$. Of special interest are the *conserved quantities*, which are the observables with zero Poisson bracket with the hamiltonian. Of course the hamiltonian itself is always a conserved quantity.

8.10 In the special case $M = R^{2n} = \{(p, q)\}$, with $\omega = dp_i \wedge dq^i$,the Poisson bracket takes the form

$$\{f, g\} = \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q^i} - \frac{\partial f}{\partial q^i}\frac{\partial g}{\partial p_i}.$$

In particular,

$$\{p_i, p_j\} = 0 = \{q^i, q^j\}, \quad \{p_i, q^j\} = \delta_i^j.$$

These are the *canonical commutation relations*. The differential equations for time evolution are

$$\frac{dq^i}{dt} = \frac{\partial H}{\partial p_i}, \quad \frac{dp_i}{dt} = -\frac{\partial H}{\partial q^i}.$$

These are *Hamilton's equations*.

8.11 In some cases these equations can be solved for arbitrary initial conditions in terms of familiar functions (such as elliptic functions). But typically the behavior of the solutions for long time intervals in extremely complicated. For example, even a small change in initial conditions can grow exponentially. The final point is often statistically independent of the starting point. And yet it is completely determined by an exact knowledge of the initial point; this peculiar phenomenon is called *chaos*. It is still poorly understood.

8.12 Mechanics is still an active and exciting branch of theoretical physics.

Chapter 9

Riemmanian Geometry

9.1 A *Riemannian metric* g on a manifold M is a symmetric non-degenerate covariant tensor field. In other words, given a pair of vector fields u and v, we get a function (their *inner product*) $g(u,v)$ which is bilinear and $g(u,v) = g(v,u)$. Moreover, $g(.,u) = 0 \Rightarrow u = 0$.

9.2 A metric g is *positive* if $g(u,u) \geq 0$ for all u. For positive metrics, the $g(u,u)$ is the square of the length of a vector. The *angle* between two vectors is $\arccos\left(\frac{g(u,v)}{\sqrt{g(u,u)}\sqrt{g(v,v)}}\right)$. If the metric is not positive, the length and angle can become complex numbers.

9.3 The simplest example is the Euclidean metric on R^n : $g(u,v) = u^i v^i$. It is clear that this metric is positive and that the length and angle as defined above have their usual meanings.

9.4 Another example is the Minkowski metric on R^4 : $g(u,v) = u^0v^0 - u^iv^i$ where $i = 1,2,3$. This describes space-time in the special theory of relativity.

9.5 Recall that a tangent vector to a sphere can be thought of as a vector in R^3 as well. Thus the Euclidean inner product on R^3 induces a metric on S^2, called the *standard metric*.

9.6 In local co-ordinates a metric is described by a set of $\frac{n(n+1)}{2}$ functions, its components: $g = g_{ij}dx^i \otimes dx^j$. Sometimes the product $\otimes$ is omitted and we just write $g = g_{ij}dx^i dx^j$. This can be viewed as the length squared of an infinitesimal arc connecting x^i to $x^i + dx^i$; thus often it is denoted by $ds^2 = g_{ij}dx^i dx^j$.

9.7 Thus in Euclidean space $ds^2 = dx^i dx^i$. The standard metric on the sphere is, in spherical polar coördinates, $ds^2 = d\theta^2 + \sin^2\theta d\phi^2$. Another interesting case is the Poincare metric on the upper half plane: $ds^2 = \frac{dx^2+dy^2}{y^2}$.

9.8 An *immersion* is a smooth map $\phi : N \to M$ whose Jacobian matrix $d\phi$ is of constant rank. An *embedding* is an immersion which is injective; i.e., its image doesnt intersect itself anywhere. Given a metric g on M , and an embdedding $\phi : N \to M$ we can get a metric (the *induced metric* $\phi^* g$)on N . Since N can be identified as a subset of M , (the image of ϕ) its tangent vectors can be thought of as tangent vectors of M ; the induced metric is just the inner product taken with g . We already saw an example of this: the metric on S^2 induced by its embedding to R^3 .

9.9 A diffeomorphism $\phi : N \to M$ is an *isometry* between Riemannian manifolds (N, h) and (M, g) if $h = \phi^* g$. Isometric manifolds are essentially the same from the point of view of Riemannian geometry.

9.10 A deep theorem of Nash and Whitney is that any Riemannian manifold with positive metric is isometric to a submanifold of some Euclidean space. This generalizes the way we obtained the metric on S^2 . In general, the dimension of the ambient Euclidean space is much higher than the dimension of the original Riemannian manifold.

9.11 Let us restrict for the rest of this section to the case of positive metrics; all the ideas can be generalized to the non-positive case as well.

9.12 The *length* of a curve $\gamma : [a, b] \to M$ is the integral of the length of its tangent vectors: $l(\gamma) = \int_a^b \sqrt{g(\dot\gamma, \dot\gamma)} dt$. It is invariant under changes of the parameter describing the curve: if $\phi : [a, b] \to [c, d]$ is a diffeomorphism of intervals, $\tilde\gamma(t) = \gamma(\phi(t))$ has the same length as γ .

9.13 The arclength $s(t) = \int_a^t \sqrt{g(\dot\gamma, \dot\gamma)} dt$ is the most natural choice of parameter along a curve. In this parametrization, the tangent vector is always of unit length so that $\sqrt{\left[g(\frac{d}{ds}\gamma, \frac{d}{ds}\gamma)\right]} ds = g(\frac{d}{ds}\gamma, \frac{d}{ds}\gamma) ds$.

9.14 A *geodesic* is curve for which the length does not change under infinitesimal variations that fix the initial and final points. We can get a differential equation for a geodesic by the Euler-Lagrange method. It is most

convenient to use the arclength as the paremeter so that we minimize the functional $\int_0^l g(\frac{d}{ds}\gamma, \frac{d}{ds}\gamma)ds$. (If we use a generic parametrization we will get the same answer, after some more calculations.)

9.15 It is best to use a co-ordinate system; i.e., consider the special case where the curve lies within one co-ordinate neighborhood and is given by some functions $x^i(t)$ for $i = 1, \cdots n$:

$$\begin{aligned}\delta \int_0^l g_{ij}(x(s))\frac{dx^i}{ds}\frac{dx^j}{ds}ds &= \int_0^l \partial_k g_{ij}\delta x^k \frac{dx^i}{ds}\frac{dx^j}{ds}ds + 2\int_0^l g_{ij}\frac{d\delta x^i}{ds}\frac{dx^j}{ds}ds \\ &= \int_0^l \partial_k g_{ij}\delta x^k \frac{dx^i}{ds}\frac{dx^j}{ds}ds - 2\int_0^l g_{ij}\delta x^i \frac{d^2x^j}{ds^2}ds \\ &- 2\int_0^l \partial_k g_{ij}\frac{dx^k}{ds}\delta x^i \frac{dx^j}{ds}ds\end{aligned}$$

9.16 The boundary term in the integration by parts can be ignored because the endponts are fixed. Relabelling indices,

$$\delta \int_0^l g_{ij}(x(s))\frac{dx^i}{ds}\frac{dx^j}{ds}ds = (-2)\int_0^l \delta x^k \left[g_{kj}\frac{d^2x^j}{ds^2} + \frac{1}{2}\left\{\partial_i g_{kj} + \partial_j g_{ki} - \partial_k g_{ij}\right\}\frac{dx^i}{ds}\frac{dx^j}{ds}\right] ds$$

9.17 Thus the differential equation for a geodesic is

$$\frac{d^2x^j}{ds^2} + \Gamma^j_{ik}\frac{dx^i}{ds}\frac{dx^k}{ds} = 0$$

where Γ^j_{ik} is a quantity called the *Christoffel symbol*:

$$\Gamma^j_{ik} = \frac{1}{2}g^{lj}\left\{\partial_i g_{kl} + \partial_i g_{kl} - \partial_l g_{ik}\right\}.$$

9.18 The geodesic equation is invariant under arbitrary changes of co-ordinates. But, the Christoffel symbols do *not* transform as tensors. For example, even if they are zero in one co-ordinate system (e.g., cartesian co-ordinates in Euclidean space) it may not be zero in a general co-ordinate system.

9.19 It follows that the geodesics of Euclidean space are straight lines; those of the sphere are the great circles. The geodesics of the Poincare metric of the upper half plane are circles that intersect with the real line at right angles.

9.20 Let us fix a point $O \in M$. There is a neighborhood U of O where each point is connected to the origin by a unique geodesic. Define a function $S : U \to R$ to be the square of the distance of the point from O . Surfaces of constant S are analogous to spheres.

9.21 The vector $g^{ij}\partial_i S \partial_j$ is orthogonal to the surface of constant S . In fact it is just the tangent vector of unit length to the geodesic connecting the point to the origin. Thus

$$g^{ij}(x)\partial_i S \partial_j S = 1.$$

This first order partial differential equation is called the *eikonal* equation. It is just another way of describing geodesics of a manifold: the geodesics are its characteristic curves. It is analogous to the Hamilton-Jacobi equation in Mechanics. Often it is easier to solve this equation than the ordinary differential equation above.

9.22 There is a natural co-ordinate system in the neighborhood of each point in a Riemannian manifold, called the *Riemann normal co-ordinates*. Choose a point $O \in M$ as the origin. There is a neighborhood U of O such that each point $P \in U$ is connected to O by a unique geodesic. Let $s(P)$ be the length of geodesic and $u(P)$ the unit tangent vector *at the origin* of the geodesic connecting O to P. Then $s(P)u(P)$ which is a vector in the tangent space to M at the origin. We can now choose an orthonormal basis in $T_O M$ and get a set of components $x^i(P)$ of this vector, which will be the normal co-ordinate of the point P . A single such co-ordinate system may not cover M ; but we can choose a number of such points spread out over the manifold with overlaping neighborhoods as usual to cover M .

9.23 The co-ordinate axes of the normal system are geodesics originating at O pointing along orthogonal directions.The system is valid as long these geodesics dont intersect. In Euclidean space, the normal co-ordinates are thus just the Cartesian co-ordinates.

Chapter 10

Parallel Transport

10.1 The components of the metric tensor $g = g_{ij}dx^i dx^j$ are constant in cartesian co-ordinates in Euclidean space. In general, there will be no such co-ordinate system . The necessary and sufficient condition for the existence of a local co-ordinate system, in which the metric tensor has constant components, is the vanishing of a quantity called the curvature or Riemann tensor.

10.2 Consider a geodesic $\gamma : [a, b] \to M$ in a two-dimensional Riemannian manifold (of positive metric). We define an operation called *parallel transport* $P(\gamma) : T_{\gamma(a)}M \to T_{\gamma(b)}M$ which takes each vector from the beginning of the curve to one at the endpoint: $P(\gamma)v$ has the same length as v and makes the same angle with the final tangent vector $\dot{\gamma}(b)$ as the v makes with the initial tangent vector $\dot{\gamma}(a)$.

10.3 If each segment of a piecewise smooth continuous curve $\gamma : [a, b] \to M$ is a geodesic, we can define the parallel transport operator on it as the product of parallel transport along each segment. Note that the parallel transport around a closed curve may not be the identity .

10.4 Any smooth curve can be approximated by geodesic segments; by choosing the segments small enough we can make the approximation as good as necessary. Thus we can define parallel transport on an arbitrary smooth curve as the limit of the transport on these approximations.

10.5 Thus a geodesic is a curve for which the tangent vector at any point is the parallel transport of the initial tangent vector.

10.6 Consider two vectors u, v at some point $p \in M$. These define a parallelogram whose sides are geodesics of infinitesimal lengths ϵ_1 and ϵ_2. The parallel transport around this parallelogram is $P \sim 1 + R(u,v)\epsilon_1\epsilon_2$. The operator $R(u,v)$ is bilinear in u and v; interchanging u, v reverses the direction in which the parallelogram is traversed, so $R(u,v) = -R(v,u)$. Thus R is a two-form valued in the vector space of linear operators on T_pM. In other words it is a tensor (called the *curvature* or *Riemman tensor*) with one contravariant index and three covariant indices.

10.7 Riemann found a generalization to higher dimensions of these ideas.

10.8 The infinitesimal change of the components of vector under parallel transport from x^i to $x^i + dx^i$ is $\delta v^i = -dx^j \Gamma^i_{jk} v^k$.

10.9 The difference between the value of the vector field at $x^i + dx^i$ and the parallel tranpsort of $v^j(x)\partial_j$ to $x^i + dx^i$ gives us the notion of the *covariant derivative*:

$$\nabla_j v^i = \frac{\partial v^i}{\partial x^j} + \Gamma^i_{jk} v^k$$

These are the components of a tensor field $\nabla v = \nabla_j v^i dx^j \otimes \partial_i$ of type $(1,1)$.

10.10 *The partial derivatives* $\frac{\partial v^i}{\partial x^j}$ *do not form the components of a tensor.* Neither do the Christoffel symbols Γ^i_{jk}. But the covariant derivative is a tensor because the inhomogenous terms in the transformation law for each term cancel out.

10.11 The parallel transport operator $P(t) : T_{\gamma(t)}M \to T_{\gamma(t)}M$ can be thought of as a matrix once we have a co-ordinate basis ∂_i on the tangent spaces. It is the solution of the ordinary differential equation

$$\frac{dP^i_j(t)}{dt} + \Gamma^i_{jk} \frac{d\gamma^k(t)}{dt} = 0$$

with the initial condition $P^i_j(0) = \delta^i_j$.

10.12 A *geodesic vector field* is one whose covariant derivative along itself is zero:

$$v^j \nabla_j v^i = 0.$$

The integral curves of a geodesic vector field are geodesics.

10.13 The partial derivatives satisfy the symmetry relation $\partial_j \partial_k v^i = \partial_k \partial_j v^i$. The covariant derivatives do not; the difference $\nabla_j \nabla_k v^i - \nabla_k \nabla_j v^i$ however does not involve a derivative of v^i :

$$\nabla_j \nabla_k v^i - \nabla_k \nabla_j v^i = \left[\partial_j \Gamma^i_{kl} - \partial_k \Gamma^i_{jl} + \Gamma^i_{jm}\Gamma^m_{kl} - \Gamma^i_{jm}\Gamma^m_{kl}\right] v^l.$$

Since the covariant derivative is a tensor, the term within square brackets must be a tensor as well. It is the *Riemann tensor* or *curvature tensor*:

$$R^i_{jkl} = \partial_j \Gamma^i_{kl} - \partial_k \Gamma^i_{jl} + \Gamma^i_{jm}\Gamma^m_{kl} - \Gamma^i_{jm}\Gamma^m_{kl}$$

10.14 There is another way to understand the curvature tensor. Given three vector fields u, v, w , imagine parallel transporting u along the integral curve of v by a distance ϵ_1 ; then along w by an amount ϵ_2 . Now do the same parallel transports but in reverse order. In general we wont arrive at the same points: they will differ by $[v, w]\epsilon_1\epsilon_2$. We do one more parallel transport along this curve to arrive at the same point. The difference between the two parallel transports is a tensor:

$$\nabla_v \nabla_w u - \nabla_v \nabla_w u - \nabla_{[v,w]} u = R(v, w)u$$

The components of this tensor $R = R^i_{jkl} dx^j \wedge dx^k \otimes dx^l \otimes \partial_i$ are as given above.

10.15 A vector field u is *covariantly constant* if

$$\nabla_v u = 0, \forall v.$$

Then

$$R(v, w)u = 0 \forall v, w.$$

Or, $R^i_{jkl} u^l = 0$. It means that the parallel transport of u along a curve depends only on the endpoint.

10.16 If and only if the curvature tensor is zero, we can find a system of n linearly independent (over R) vector fields in a small enough neighborhood of any point. The metric tensor is then constant in the Riemann normal co-ordinates. Thus the curvature tensor measures locally the departure from Euclidean geometry.

10.17 The simplest example of a Riemannian manifold of non-zero curvature is a sphere.

Chapter 11

Connection and Curvature

11.1 We can define the notion of covariant derivative in an axiomatic way, analogous to the way we introduced the exterior derivative.

11.2 A *covariant derivative* or *connection* on a differential manifold is a map $\nabla : \mathcal{T}^r_s(M) \to \mathcal{T}^r_{s+1}(M)$ such that

$$\begin{aligned} \nabla(S \otimes T) &= \nabla S \otimes T + S \otimes \nabla T && (11.1)\\ \nabla f &= df, \forall f \in \mathcal{T}^0_0(M) = C(M) && (11.2)\\ d\left[\phi(u)\right] &= \nabla\phi(u) + \phi\left[\nabla u\right] && (11.3) \end{aligned}$$

11.3 We can translate these into tensor index notation

$$\begin{aligned} \nabla_m(S^{i_1\cdots i_p}_{j_1\cdots j_q}T^{k_1\cdots k_r}_{l_1\cdots l_s}) &= \left[\nabla_m S^{i_1\cdots i_p}_{j_1\cdots j_q}\right]T^{k_1\cdots k_r}_{l_1\cdots l_s} + S^{i_1\cdots i_p}_{j_1\cdots j_q}\left[\nabla_m T^{k_1\cdots k_r}_{l_1\cdots l_s}\right] && (11.4)\\ \nabla_i f &= \partial_i f, \forall f \in \mathcal{T}^0_0(M) = C(M) && (11.5)\\ \partial_i\left[\phi_j u^j\right] &= \left[\nabla_i\phi_j\right]u^j + \phi_j\left[\nabla_i u^j\right] && (11.6) \end{aligned}$$

11.4 The covariant derivative of a vectorfield satisfies as a consequence

$$\nabla_i\left[fu^j\right] = \partial_i f u^j + f\nabla_i u^j, \forall f \in C(M).$$

It follows that the covariant derivative must be involve at most first derivatives of the components of u :

$$\nabla_i u^j = \partial_i u^j + \Gamma^j_{ik}u^k$$

for some some set of quantities Γ^j_{ik} called the *connection coefficients*.

11.5 The covariant derivatives on one-forms has a similar form. The third axiom now determines the connection coefficients of the one-forms in terms of the very same Γ :

$$\nabla_i \phi_j = \partial_i \phi_j - \Gamma^k_{ij} \phi_k.$$

11.6 The first axiom now gives the general formula for covariant derivative (a tensor is a linear combination of tensor products of vector fields and one-forms)

$$\begin{aligned} \nabla_k T^{i_1 \cdots i_p}_{j_1 \cdots j_q} = & \\ & \nabla_k T^{i_1 \cdots i_p}_{j_1 \cdots j_q} + \\ \Gamma^{i_1}_{kl} T^{l i_2 \cdots i_p}_{j_1 \cdots j_q} \quad + \quad & \Gamma^{i_2}_{kl} T^{i_1 l i_3 \cdots i_p}_{j_1 \cdots j_q} + \cdots \Gamma^{i_p}_{kl} T^{i_1 i_2 \cdots l}_{j_1 \cdots j_q} \\ -\Gamma^{l}_{k j_1} T^{i_1 i_2 \cdots i_p}_{l j_2 \cdots j_q} \quad - \quad & \Gamma^{l}_{k j_2} T^{i_1 i_2 \cdots i_p}_{j_1 l j_3 \cdots j_q} - \cdots \Gamma^{l}_{k j_q} T^{i_1 i_2 \cdots i_p}_{j_1 j_2 \cdots l}. \end{aligned} \tag{11.7}$$

11.7 The covariant derivative along a vector u of a tensor is defined by contracting it with the connection:

$$\nabla_v T^{i_1 \cdots i_p}_{j_1 \cdots j_q} = v^k \nabla_k T^{i_1 \cdots i_p}_{j_1 \cdots j_q}$$

11.8 The parallel transport operator $P(t) : T_{\gamma(a)} M \to T_{\gamma(t)} M$ along a curve $\gamma : [a, b] \to M$ is given by the solution of the ordinary differential equation

$$\frac{dP^i_j(t)}{dt} + \Gamma^i_{jk} \frac{d\gamma^k(t)}{dt} = 0$$

with the initial condition

$$P^i_j(a) = \delta^i_j.$$

The parallel transport depends on the curve chosen in general, not just on the endpoints.

11.9 The anti-symmetric part of the second covariant derivative on functions defines a tensor called *Torsion*

$$\nabla_i \nabla_j f - \nabla_i \nabla_j f = T^k_{ij} \partial_k f, \quad T^k_{ij} = \Gamma^k_{ij} - \Gamma^k_{ji}.$$

Although Γ^k_{ij} do not transform as the components of a tensor, its antisymmetric part does.

11.10 We will only consider connections for which the torsion tensor vanishes. This will be an extra axiom:

$$\nabla_u \nabla_v f - \nabla_v \nabla_u f - \nabla_{[u,v]} f = 0, \forall f \in C(M).$$

11.11 Given a Riemann metric $g = g_{ij} dx^i \otimes dx^j$ it is natural to restrict ourselves to connections that preserve the inner prodcuts of vector fields. This is the axiom

$$\nabla g = 0$$

11.12 There is a unique connection on a Riemann manifold (M, g) satisfying the above axioms. To see this we use the symmetry of the connection coefficients $\Gamma^i_{jk} = \Gamma^i_{kj}$ and the equation

$$\nabla_i g_{jk} = \partial_i g_{jk} - \Gamma^l_{ij} g_{lk} - \Gamma^l_{ik} g_{jl} = 0$$

to solve for the Γ^i_{jk} :

$$g_{il} \Gamma^l_{jk} = \frac{1}{2} \left[\partial_k g_{ij} + \partial_j g_{ik} - \partial_i g_{jk} \right].$$

11.13 The difference of covariant derivatives taken in two different order gives the *curvature tensor* or *Riemann tensor*:

$$\nabla_u \nabla_v w - \nabla_v \nabla_u w - \nabla_{[u,v]} w = R(u, v) w.$$

It is a non-trivial calculation to check that this is indeed a tensor:

$$R(fu, v)w = fR(u, v)w, \quad R(u, v)[fw] = fR(u, v)w.$$

This can also be thought of the difference between parallel transports along two different infinitesimally small curves, with sides $u\epsilon_1, v\epsilon_2$ and $v\epsilon_2, u\epsilon_1, [u, v]\epsilon_1\epsilon_2$ respectively.

11.14 In terms of a co-ordinates basis, we have $R = R^k_{ijl} dx^i \otimes dx^j \otimes \partial_k \otimes dx^l$.

$$R^k_{ijl} = \partial_i \Gamma^k_{jl} - \partial_j \Gamma^k_{il} + \Gamma^k_{im}\Gamma^m_{jl} - \Gamma^k_{im}\Gamma^m_{jl}$$

The curvature tensor satisfies the identities

$$R^k_{ijl} = -R^k_{jil}, \quad R^k_{ijl} + R^k_{jli} + R^k_{lij} = 0.$$

as well as the *Bianchi identity*

$$\nabla_m R^i_{jkl} + \nabla_k R^i_{mjl} + \nabla_j R^i_{kml} = 0.$$

11.15 The parallel transport operator preserves the length of a vector. The curvature tensor, being the infinitesimal parallel transport operator around a closed curve, must preserve the length of a vector:

$$g(R(v,w)u, r) + g(u, R(v,w)r) = 0.$$

If we define $g_{km} R^m_{ijl} = R_{ijkl}$, we have thus $R_{ijkl} = -R_{ijlk}$.

11.16 Finally, we can show that $R_{ijkl} = R_{klij}$.

11.17 From the Riemann tensor we can construct some other standard objects of Riemannian geometry. The *Ricci tensor* is

$$R^j_{ijl} = R_{il}$$

and its trace is the *Ricci scalar*

$$R = g^{jk} R_{jk}.$$

11.18 A Riemannian manifold is *flat* if there is a co-ordinate system such that the metric tensor has constant components. The basic theorem of Riemannian geometry is that *a Riemannian manifold is flat if and only if the curvature tensor vanishes everywhere.* In fact the metric tensor has constant components in Riemann normal co-ordinates.

Chapter 12

Geometrical Optics

12.1 When the wavelength of light is small, it is possible to approximate its propagation by rays. This is the limit of *geometrical optics*.

12.2 The speed v of propagation of light in a medium is less than that in vacuum, c. The ratio $\frac{c}{v} = n$ is called the *refractive index*. It can depend on position. For glass it varies in the range $n \sim 1.3 - 1.8$.

12.3 Given a curve $\mathbf{r} : [a, b] \to R^3$ in space, its *optical length* is defined to be $\int_a^b n(\mathbf{r}(t))|\dot{\mathbf{r}}(t)|dt$. This is the length of the curve with respect to the *optical metric* $ds_o^2 = n(x)^2[dx^2 + dy^2 + dz^2]$. This is a Riemann metric on R^3 which can differ from the Euclidean metric by the *conformal factor* $n(x)$. The angles between vectors with respect to the optical metric are the same as in Euclidean geometry, but the lengths of vectors (and hence curves) are different.

12.4 The basic law of geometrical optics is *Fermat's principle*: *light propagates along curves of extremal optical length*. In other words *light propagates along the geodesics of the optical metric*.

12.5 If we parametrize the curve $\mathbf{r}(s)$ by the *Euclidean arc length* s, the tangent vector of unit $\dot{\mathbf{r}}(s)$ *Euclidean length*. Then the *optical length* of the curve is given by $\int n(\mathbf{r}(s))\dot{\mathbf{r}}(s)^2 ds$; it is an extremum when the ordinary differential equation

$$\frac{d}{ds}\left(n\frac{d\mathbf{r}}{ds}\right) = \mathbf{grad}\ n$$

is satisfied.

12.6 Given a function $S : R^3 \to R$, we have a finally of surfaces on which S is a constant. Given the Euclidean metric on R^3 , we have a family of curves (*characteristic curves*), which are orthogonal to these surfaces.

12.7 If we choose S to be the optical distance from a fixed point (say the origin) the characteristic curves are the light rays. Since $\frac{d\mathbf{r}}{ds} \cdot \mathbf{grad}\ S = n$ is the difference between the optical distances between two points, we have

$$(\mathbf{grad}\ S)^2 = n^2.$$

This *eikonal equation* contains the same information as the geodesic equation. The surfaces of constant S are called *wavefronts*.

12.8 The fundamental physical laws of optics are the Maxwell equations, since light is just a kind of electromagnetic field. It should be possible to derive the eikonal equation and the Fermat principle from the Maxwell equations,

$$\mathbf{curl}\ \mathbf{H} = \frac{1}{c}\dot{\mathbf{D}}, \quad \mathbf{curl}\ \mathbf{E} = -\frac{1}{c}\dot{\mathbf{B}}$$

$$\mathbf{div}\ \mathbf{E} = 0, \quad \mathbf{div}\ \mathbf{B} = 0.$$

In addition we have the *constitutive relations*

$$\mathbf{D} = \epsilon\mathbf{E}, \quad \mathbf{B} = \mu\mathbf{H}(x).$$

Here we have assumed that the medium is time static and isotropic and that there are no electric charges or currents . The function ϵ is *dielectric permittivity* and μ is the *magnetic permeability*.

12.9 If ϵ and μ are constants, there are plane wave solutions: $\mathrm{E}(x,t) = \mathrm{Re}\ \mathbf{e}\ e^{i\mathrm{k}\cdot\mathrm{x}-\omega t}, \mathrm{H}(x,t) = \mathrm{Re}\ \mathbf{h}\ e^{i\mathrm{k}\cdot\mathrm{x}-\omega t}$. The phase of the wave is constant along the planes $\mathbf{k}\cdot\mathbf{x} = \text{constant}$.

12.10 We will assume that ϵ and μ are slowly varying functions: i.e., over one wavelength they change by a small amount. In this case, there should be a modification of the plane wave solution, where the wavefronts are some curved surfaces, and the amplitudes $\mathbf{e}(x)$ and $\mathbf{h}(x)$ are slowly varying functions of x :

$$\mathrm{E}(x,t) = \mathrm{Re}\ \mathbf{e}(x)\ e^{ik_0 S(x)-\omega t}, \quad \mathrm{H}(x,t) = \mathrm{Re}\ \mathbf{h}\ e^{ik_0 S(x)-\omega t}$$

Here, $k_0 = \frac{\omega}{c}$. The wavefronts are the surfaces of constant S .

12.11 We now put this into Maxwell's equations and ignore derivatives of $\epsilon, \mu, \mathbf{e}, \mathbf{h}$ to get

$$\mathbf{grad}\ S \times \mathbf{h} + \epsilon \mathbf{e} = 0, \quad \mathbf{grad}\ \times \mathbf{e} - \mu \mathbf{h} = 0.$$

The remaining equations follow from these.

12.12 This system of algebraic equations have a solution only if

$$(\mathbf{grad}\ S)^2 = \epsilon\mu$$

as we can see by eliminating $\mathbf{h}$ and using $\mathbf{e} \cdot \mathbf{grad}\ S = 0$. If we identify $n = \sqrt{[}\epsilon\mu]$ with the refractive index, we get the eikonal equation.

12.13 How do the amplitudes $\mathbf{e}$ and $\mathbf{h}$ change as a wave propagates along the geodesics?. Clearly they are orthogonal to the direction of propagation $\mathbf{grad}\ S$ and to each other. From Maxwell's equations we can derive in a similar way

$$\frac{\partial}{\partial \tau}\mathbf{e} + [\mathbf{e} \cdot \mathbf{grad}\ \ln n]\mathbf{grad}\ S = 0, \quad \frac{\partial}{\partial \tau}\mathbf{h} + [\mathbf{h} \cdot \mathbf{grad}\ \ln n]\mathbf{grad}\ S = 0$$

where $\frac{\partial}{\partial \tau} = \mathbf{grad}\ S \cdot \mathbf{grad}$ is the derivative along the light ray. These are precisely the equations of parallel transport with respect to the optical metric.

12.14 Thus if two points are connected by two different optical geodesics, there is not only a phase difference between the two paths due to their different optical lengths, but also a relative rotation of polarization due to the parallel transport. This rotation of the polarization (often called the *Pancharatnam*, *geometric* or *Berry phase*) can be observed easily using laser beams propagating along two different optical fibers.

12.15 Maxwell proposed an optical instrument called the *fish-eye*. It has the refractive index

$$n(x) = \frac{n_0}{1 + \frac{\mathbf{x}^2}{a^2}}.$$

It is an example of an *absolute* optical instrument; i.e., from every point outside the circle of radius a , there are an infinite number of rays that

connect it to an image point inside the circle. (This image point is the inversion with respect to the circle of the original point.) The rays are circles orthogonal to the circle of radius a . These facts follow easily once we realize that the optical metric corresponding to the fish-eye is the *standard metric on a three-dimensional sphere* of radius a :

$$ds_o^2 = \frac{1}{1 + \frac{\mathbf{x}^2}{a^2}}[dx^2 + dy^2 + dz^2].$$

Chapter 13

Special Relativity

13.1 It was discovered at the end of the last century that the velocity of light c in the vacuum is the same for all observers. This was inconsistent with the prevailing view of the time that light is the oscillation of a medium (called ether) that pervaded all matter, just as sound is the oscillatio of air.

13.2 It was also inconsistent with the laws of newtonian mechanics, but was consistent with the laws of electromagnetism (Maxwell's equations). Einstein realized that the laws of mechanics can be changed to fit the new theory of electromagnetism: the departure from newtonian mechanics would be experimentally measurable at velocitie comparable to c .

13.3 For example, the formula for the kinetic energy and momentum of a particle of mass m are

$$E = \frac{mc^2|\mathbf{v}|}{\sqrt{\left[1 - \frac{\mathbf{v}2}{c^2}\right]}}, \quad \mathbf{p} = \frac{m\mathbf{v}}{\sqrt{\left[1 - \frac{\mathbf{v}2}{c^2}\right]}}.$$

In other words the newtonian relation $E = \frac{\mathbf{p}^2}{2m}$ is replaced by

$$E^2 - \mathbf{p}^2c^2 = m^2c^4.$$

13.4 Moreover, the formulas of transformation between two observers moving at a constant velocity with respect to each other are changed from the *Galilean transformations*

$$\mathbf{x}' = \mathbf{x} - \mathbf{v}t, \quad t' = t$$

to the *Lorentz transformations*

$$\mathbf{x}' = \frac{\mathbf{x} - \mathbf{v}t}{\sqrt{\left[1 - \frac{\mathbf{v}^2}{c^2}\right]}}, \quad t' = \frac{t - \frac{1}{c^2}\mathbf{v}\cdot\mathbf{x}}{\sqrt{\left[1 - \frac{\mathbf{v}^2}{c^2}\right]}}.$$

13.5 In particular, this implies that times measured by different observers are not the same. Two observers can even disagree on whether two events occuring at different points in space are simultaneous.

13.6 Minkowski realized that this *theory of relativity* had a simple meaning in geometric terms. He postulated that space and time be combined into a four dimensional manifold with the metric

$$ds^2 = \mathbf{dx}^2 - c^2dt^2.$$

This *Minkowski metric* is not positive: vectors of positive length^2 are said to be *space-like*; those of negative length^2 *time–like*; there are non-zero vectors of zero length which are said *null* or *light–like*.

13.7 A basic fact of physics is causality: *it is possible to send a signal from one event only to its future.* This principle continues to hold in relativity with the understanding that *an event* x *is in the future of* y *if* $\eta(x-y, x-y) \leq 0$ *and* $(x^0 - y^0) > 0$. The condition to be in the future is independent of changes of reference frames under Lorentz transformations only if the first condition is also satisfied: the points cannot be space-like separated.

13.8 If they are space-like separate (i.e., if $\eta(x - y, x - y) > 0$) there are reference frames in which $x^0 - y^0$ positive, negative or zero. Thus the principle of causality implies that *it is not possible to send a signal between two events that are space-like separated.* What this means is that *no signal can travel at a velocity greater than that of light.*

13.9 The motion of a particle is given by a curve in space-time called its *world-line*. The principle of causality implies that the tangent vector to the world-line of a particle is either time-like or null.

13.10 Define the length of a time-like worldline (i.e., one whose tangent vector is time–like everywhere) $x : [a, b] \to R^4$ to be

$$S[x] = \int_a^b \left[-\eta(\frac{dx}{d\tau}, \frac{dx}{d\tau}) \right]^{\frac{1}{2}} d\tau.$$

13.11 The length of a time-like curve can be thought of as the time as measured by an observer moving with along the curve, multiplied by c . It is also called the *proper time* of the curve. It is invariant under the changes of the parametrization; i.e., two curves x and $\tilde{x}$ have the same length if there is a monotonic function $\phi : [a, b] \to [a, b]$ preserving endpoints such that

$$\tilde{x}(\tau) = x(\phi(\tau)).$$

Still, the most natural parameter to use in describing the curve is the arc-length, given by the indefinite integral

$$s(\tau) = \int_a^\tau \left[-\eta(\frac{dx}{d\tau}, \frac{dx}{d\tau}) \right]^{\frac{1}{2}} d\tau.$$

The tangent vector with respect to the arc-length is of unit length:

$$\eta(\frac{dx}{ds}, \frac{dx}{ds}) = -1.$$

13.12 *Free particles move along curves for which the arc–length is an extremum.* In other words *free particles move along time–like geodesics of Minkowski metric.* It is easy to see that these are straight-lines with time-like tangents. Thus this principle is the relativistic analogue of Newton's first law.

13.13 The unit tangent vector to the world-line $\frac{v^\mu = dx^\mu}{ds}$ is often called the *four-velocity* of the particle. It is important to remember that its spatial components are *not* the velocity $\mathbf{u} = \frac{d\mathbf{x}}{dt}$ as defined in Newtonian mechanics. Instead,

$$\mathbf{v} = \frac{\mathbf{u}}{\sqrt{[1 - \frac{\mathbf{u}^2}{c^2}]}}.$$

13.14 The relatistic formulae relating momentum and energy to velocity now becomes

$$p_\mu = mc\eta_{\mu\nu}\frac{dx^\nu}{ds};$$

The relation between energy and momentum becomes

$$\eta^{\mu\nu}p_\mu p_\nu = -(mc)^2.$$

13.15 The variational equations that describe a geodesic in arbitrary parameters makes sense even when the tangent is not time–like. Since the length of the tangent is constant, if it is null at one point it will be null everywhere. Null geodesics are defined as the solutions of this equation with null tangents. They are just straight-lines with null tangents. *Light travels along null geodesics*. This is the relativistic analogues of Fermat's principle. The arclength of a null-geodesic is zero, so cannot be used as a parameter on the curve.

Chapter 14

Electrodynamics

14.1 The equations that determine the electric and magnetic fields in vacuum with electric charge density ρ and current density $\mathbf{j}$ are **Maxwell's equations**:

$$\mathbf{div}\ B = 0, \quad \mathbf{curl}\ E = -\frac{1}{c}\frac{\partial \mathbf{B}}{\partial t}$$

and

$$\mathbf{div}\ E = 4\pi\rho, \quad \mathbf{curl}\ B = \frac{4\pi}{c}\mathbf{j} + \frac{1}{c}\frac{\partial \mathbf{E}}{\partial t}.$$

14.2 It is a consequence that the electric charge is conserved:

$$\frac{1}{c}\frac{\partial \rho}{\partial t} - \mathbf{div}\ \mathbf{j} = 0.$$

14.3 A charge distribution that is static (i.e., with $\mathbf{j}=\mathbf{0}$) to one observer will look moving to another. Under Lorentz transformations, ρ and $\mathbf{j}$ transform into each other: they together transform as a four vector $j = (\rho, \mathbf{j})$. The conservation of charge then is the partial differential equation

$$\partial_\mu j^\mu = 0.$$

14.4 Electric and magnetic fields are transformed into each other as well under Lorentz transformations. Together they must transform as a tensor with six independent components. The simplest possibility is that they form a two-form:

$$F = F_{\mu\nu} dx^\mu \wedge dx^\nu, \quad F_{0i} = E_i, \quad F_{12} = B_3, F_{23} = B_1, F_{31} = B_2.$$

This is indeed the correct trasnformation law.

14.5 The pair Maxwell equations that do not involve sources can now be combined into a single equation that is Lorentz invariant:

$$dF = 0.$$

This means within any local neighborhood there is a one-form A such that $F = dA$. In other words, with $A = A_0 dx^0 + A_i dx^i$

$$\mathbf{E} = \frac{\partial \mathbf{A}}{c\partial t} - \mathbf{grad}\ A_0, \quad \mathbf{B} = \mathbf{curl}\ \mathbf{A}.$$

Thus A_0 is the *electrostatic potential* and $\mathbf{A}$ the **vectorpotential** .

14.6 Note that if we add the deirivative of a scalar to A the electromagnetic field is invariant:

$$A \mapsto A + d\lambda, F \mapsto F.$$

Thus only functions of A that are invariant under this *gauge transformation* are of physical significance.

14.7 The remanining Maxwell's equations can be written as

$$\partial^\mu F_{\mu\nu} = j_\nu.$$

The conservation of current is an immediate consequence of the anti-symmetry of $F_{\mu\nu}$.

14.8 We can derive the aove equation from a variational principle, where A is thought as the fundamental variable. The only gauge and Lorentz invariant quantities that can be formed from the first derivatives of A are

14.9 $\int F_{\mu\nu}F_{\rho\sigma}\eta^{\mu\rho}\eta^{\nu\sigma}d^4x$ and $\int F\wedge F$. The latter can be excluded because it is a total derivative, $F\wedge F = d(A\wedge F)$: it would lead to trivial equations of motion. The combination $\int j_\mu A_\mu d^4x$ is gauge invariant when the current is conserved. The Maxwell equations are the Euler-Lagrange equations of the action

$$S = -\frac{1}{4}\int F_{\mu\nu}F_{\rho\sigma}\eta^{\mu\rho}\eta^{\nu\sigma}d^4x + \int j^\mu A_\mu d^4x.$$

14.10 If we have a particle moving along a worldline $\xi(\tau)$, the quantity $\int_\xi A = \int A_\mu(\xi(\tau))\frac{d\xi^\mu}{d\tau}d\tau$ is gauge invariant. The action principle

$$m\int\sqrt{\left[\eta_{\mu\nu}\dot\xi^\mu\dot\xi^\nu)\right]}\,d\tau + e\int A_\mu(\xi(\tau))\frac{d\xi^\mu}{d\tau}d\tau$$

is thus gauge, Lorentz and reparametrization invariant. It leads to the equation of motion

$$m\frac{d^2x^\mu}{ds^2} = eF^\mu{}_\nu\frac{d\xi^\nu}{ds}$$

where s is the proper-time parameter. The quantity on the r.h.s. is the force on a particle due to an electromagnetic field.

14.11 In the above we have ignored the fact that an accelerating charge will radiate. This in turn will change its acceleration: there is a recoil due to the momentum carried away by the radiation. Dirac derived a remarkable equation for the motion of a charged particle including these effects. It is a third order non-linear equation. See the book by A. O. Barut for details.

Chapter 15

General Relativity

15.1 *General Relativity* is the most accurate theory of gravity, due to Albert Einstein.

15.2 We saw that in special relativity, space-time is described as a flat Riemannian manifold of signature $-+++$. The co-ordinate systems in this manifold correspond physically to reference frames. The special co-ordinate systems in which the metric has the constant form

$$ds^2 = -c^2dt^2 + (dx^1)^2 + (dx^2)^2 + (dx^3)^2$$

correspond to *inertial reference frames*; i.e., those that move at a constant velocity.

15.3 However, in geometry we do not allow any special significance to a particular class of co-ordinate systems; we therefore seek a theory of relativity which is valid in any reference frame, even accelerating ones. That is, we postulate the *principle of general relativity*:

The laws of Physics are the same in all reference frames.

15.4 Such a theory must necesarily include a description of gravity. For, we also have the *Principle of Equivalence*, justified by direct observation:

All test-particles have the same acceleration due to gravity, regardless of their composition.

Here, a test-particle is one size and mass are small compared to the size and mass of the body creating the gravitational field: for example an artifical satellite would be a test-particle for the Earth's gravitational field; the earth itself is a test–particle for the sun; the sun is one for the cosmological gravitational field due to all the galaxies in the universe and so on.

15.5 Thus by local meausrements we cannot distinguish between a gravitational field and a reference which is accelerating at a constant rate: a uniform gravitational field is just flat space-time viewed in a non-inertial co-ordinate system.

15.6 We have not succeeded yet in producing a quantum theory that satisfies the above principles (perhaps they need to be modified.) Einstein produced a theory of gravity that satisfied both of these principles.

15.7 The first postulate of this theory is

Axiom 15.1 Space-time is a four-dimensional Riemannian manifold, whose metric g is of signature $-+++$.

If the metric is the Minkowski metric on R^4 , there is no gravitational field. Thus the curvature of the metric is a measure of the strength of the gravitational field.

15.8 The next axiom satisfies the equivalence principle since it treats all test-particles the same way, regardless of constitution:

Axiom 15.2 The metric tensor g of space-time describes the gravitational field; in the absence of forces other than gravity, test-particles move along time-like geodesics of g .

Thus if the path of a particle is described by a function $x^\mu(s)$ of arc-length s , it satisfies the differential equation

$$\frac{d^2x^\mu(s)}{ds^2} + \Gamma^\mu_{\nu\rho}(x(s))\frac{dx^\nu(s)}{ds}\frac{dx^\rho(s)}{ds} = 0.$$

15.9 Thus in Minkowski space particles will move along time-like straight lines.

15.10 In the limit when the velocities of particles are small compared to the velocity of light, this should reduce to the newtonian equations of motion

$$\frac{d^2\mathbf{x}(t)}{dt^2} + \mathbf{grad}\ \phi(x(t)) = 0$$

where ϕ is the gravitational potential. In this limit,

$$s \sim ct = x^0, \quad \frac{dx^0}{ds} \sim 1, \quad \frac{d\mathbf{x}}{ds} \sim \frac{d\mathbf{x}}{cdt}$$

so that we recover the newtonian limit if we identify

$$g_{00} \sim 1 + c^{-2}\phi.$$

15.11 The source of the gravitational field in the newtonian limit is mass-density ρ :

$$-\nabla^2\phi = 4\pi G\rho$$

where G is Newton's constant of gravitation.

15.12 Already in special relativity, mass and energy are interchangeable: they are not separately conserved.

15.13 Moreover, energy has to combined with momentum to obtain a Lorentz co-variant notion, that of four-momentum. The density of energy-momentum is the *stress-energy* energy tensor $T_{\mu\nu}$: the mass-energy density is $\rho = T_{00}$ and momentum density is T_{0i} for $i = 1, 2, 3$. The components T_{ij} describe the stress: the pressure exterted by matter in the direction i on a surface whose normal points along the co-ordinate direction j . The conservation of energy-momentum requires this tensor to satisfy

$$\partial^\mu T_{\mu\nu} = 0$$

in special relativity (i.e., in flat space-time). Moreover, conservation of angular momentum requires it to be symmetric:

$$T_{\mu\nu} = T_{\nu\mu}.$$

15.14 In general relativity, matter should described a symmetric tensor satisfying the covariant conservation law

$$\nabla^\mu T_{\mu\nu} = 0.$$

15.15 The gravitationl field must be described by an equation of the form $G_{\mu\nu} = 4\pi G T_{\mu\nu}$, where $G_{\mu\nu}$ must be a tensor built from the metric tensor which is (i) symmetric,, (ii) satisfies the identity $\nabla^\mu G_{\mu\nu} = 0$, (iii) must involve at most second derivatives of the components of the metric tensor. The latter requirement is suggested by the fact that the newtonian equation for the gravitational potential is a second order partial differential equation.

15.16 The curvature tensor $R^\mu_{\nu\rho\ \sigma}$ involved second derivatives of the metric tensor. From it we can construct the symmetric tensor $R_{\mu\nu} = R^\rho_{\mu\rho\ \nu}$ called the *Ricci tensor*; any other contraction will produce the same tensor upto a sign or zero. Further contraction yields the *Ricci scalar*, $R = g^{\mu\nu} R_{\mu\nu}$. Thus it is reasonable to suppose that $G_{mu\nu}$ is a linear combination $G_{\mu\nu} = aR_{\mu\nu} + bg_{\mu\nu}R$. The Bianchi identity on the curvature tensor gives $2\nabla^\mu R_{\mu\nu} = \partial_\mu R$. Thus a possible choice satisfying the identity $\nabla^\mu G_{\mu\nu} = 0$ is

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R.$$

This is called the *Einstein tensor*.

Axiom 15.3 The field equations of general relativity that determine the metric tensor are $G_{\mu\nu} = 8\pi G T_{\mu\nu}$.

15.17 It is possible to modify the above equations while retaining the property that they are second order partial differential equations:

$$G_{\mu\nu} - \Lambda g_{\mu\nu} = 8\pi G T_{\mu\nu}.$$

This modification of the theory depends on the *cosmological constant* Λ ; (barring some recent unconfirmed observations) this parameter is found to be zero and we will from now on let it vanish.

15.18 Einsteins equations also follow from a variational principle, due to Hilbert:

$$S = \int \sqrt{[-\det g]} R d^4x + 8\pi G S_m$$

where S_m is the action of the matter field. The point is that the variational derivative of the first term is the Einstein tensor and that of the second term the stress-energy tensor.

Chapter 16

Cohomology

16.1 Recall that the exterior derivative $d : \Lambda^p(M) \to \Lambda^{p+1}(M)$ satisfies

$$d(d\omega) = 0$$

for any ω . A form ϕ is *closed* if $d\phi = 0$; it is *exact* if there is a form ω such that $\phi = d\omega$. Clearly all exact forms are closed.

16.2 Let $Z^p(M)$ be the space of closed differential forms, and $B^p(M)$ the subspace of exact forms. The quotient space $H^p(M) = Z^p(M)/B^p(M)$ is called the *cohomology* of the manifold. For compact manifolds $\dim H^p(M)$ is finite dimensional even though $Z^p(M)$ and $B^p(M)$ can both be infnite dimensional. The number $b^p(M) = \dim H^p(M)$ is called the p -th Betti number of the manifold. It depends only on the homotopy type of M .

16.3 For example, let $M = S^1$. The space of functions on a circle can be thought of as the space of smooth functions on the real line which are periodic of period 2π . More generally, a differential form on S^1 is a form on R that is invariant under translations by 2π . Thus the co-ordinate $0 \leq \theta < 2\pi$ itself is not a function on S^1 but $\omega = d\theta$ is a differential form on the circle : $d(\theta) = d(\theta + 2\pi)$. Now, $d\omega = 0$; indeed all 1-forms on S^1 are closed; however it is not exact. The most general 1-form is a periodic function times ω : $f(\theta)\omega$ which is closed for any f . It is exact if and only if the indefinite integral of f is a periodic function. That is, iff

$$\int_0^{2\pi} f(\theta)d\theta = 0.$$

Any 1-form can be written as the sum of an exact 1-form and a multiple of ω :

$$f(\theta)\omega = [f - \bar{f}]\omega + \bar{f}\omega$$

where $\bar{f} = \frac{1}{2\pi}\int_0^{2\pi} f(\theta)d\theta$ is the mean value of f . Thus $\dim H^1(S^1) = 1$. Similarly, $\dim H^0(S^1) = 1$

16.4 More generally,

$$\dim H^p(S^n) = 1; \text{for} p = 0, n$$

and zero otherwise.

16.5 Recall that d is a graded derivation with respect to exterior product: $d(\omega \wedge \phi) = d\omega \wedge \phi + (-1)^{\deg \omega} \wedge d\phi$. So the product of two closed forms is closed. Also, $\omega \wedge d\phi = d[\omega \wedge \phi]$ if $d\omega = 0$; i.e, the product of a closed form with an exact form is exact. Stated differently, $\oplus_p B^p(M)$ is an ideal in the associative algebra $\oplus_p Z^p(M)$. Thus the quotient $\oplus_p H^p(M)$ is an associative (graded commutative) algebra, under the wedge product.

16.6 Let us give some examples. The *Grassmannian* $Gr_k(n)$ is the set of all k -dimensional subspaces of the complex vector space C^n . It is a compact manifold; also, $\dim Gr_k(n) = 2k(n-k)$. (The special case $k = 1$ is the space of rays in n and is also called the *Complex Projective space* CP^{n-1} .) We can embed the Grassmannian into a Euclidean space, the space of hermitean $n \times n$ matrices of square one and having k negative eigenvalues :

$$Gr_k(n) = \{\Phi^\dagger = \Phi; \Phi^2 = 1;\ \mathrm{tr}\Phi = n - 2k\}.$$

To each such matrix we can associate a subspace of dimension k ; the eignespace with eigenvalue -1 . Conversely given a subspace of dimension k , there is exactly one hermitean matrix with eigenvalue -1 on this subspace and $+1$ on the orthogonal complement. Grassmannians are interesting examples of differential manifolds; they are simple enough to be studied in detail yet complicated enough to serve as 'building blocks' of cohomology theory.

16.7 The differential forms $\omega_{2r} = \mathrm{tr}\Phi(d\Phi)^{2r}$ are closed: $d\omega_{2r} = \mathrm{tr}[d\Phi]^{2r+1} = 0$ since Φ anticommutes with $d\Phi$. These ω_{2r} generate the cohomology algebra of the Grassmannian. There are some relations among them, since the wedge of too high an order should vanish. In the inductive limit $n \to \infty$ these relations dissappear and the cohomology algebra of the Grassmannian $Gr_k(\infty)$ is thus a free commutative algebra generated by ω_{2r} for $r = 1, \cdots \infty$. This is related to the theory of characteristic classes (see the book *Characteristic Classes* by J. Milnor.)

Chapter 17

Morse Theory

17.1 In physics and geometry the problem of finding the extrema of a function arise often. For example the time evolution of a classical system is given by the extremum of the action; its static solutions are the extrema of potential energy; a geodesic is an extremum of the distance; the path of a light ray is the extremum of the optical distance.

17.2 Morse theory examines the question of whether a smooth function on a given manifold must have an extremum, even if the function does not have any special symmetries.

17.3 A point $p \in M$ is a *critical point* for a function $f : M \to R$ if $df(p) = 0$; thus a critical point is the same as an extremum. A critical point p is *isolated* if the second derivative $\frac{\partial^2 f}{\partial x^i \partial x^i}$ is a non-degenerate matrix at p . This property is independent of the choice of co-ordinates, at a critical point. For, the matrix of second derivatives $\partial^2 f$ transforms homogenously under co-ordinate transformations: $\partial^2 f \to S \partial^2 f S^T$, where S the matrix of first derivatives of the transformation at p . The matrix $\partial^2 f$ is called the *hessian*.

17.4 A smooth function $f : M \to R$ is a *Morse function* if all its critical points are isolated.

17.5 In the vector space of all smooth functions, the Morse functions form an open subset; the complement of this set has non-zero co-dimension; thus the space of Morse functions is like the set of vectors in R^3 with non-zero component along the x -axis. Thus a Morse function is 'generic'. Other

examples of generic subsets are knots (one-dimensional submanifolds in the space of all curves in a 3-manifolds); plane curves that intersect transversally; the orbits on a chaotic system that are not periodic.

17.6 The hessian matrix of a function at a critical point can be brought to the form $\text{diag}\ (+\,+\,+\cdots+\,-\,-\,-\cdots)$ by a choice of co-ordinates. Thus the only co-ordinate invariant information in the Hessian is the number of negative eigenvalues, which is called the *index* of a critical point. If M is compact, any smooth function f on it has a finite number $C_p(f)$ of critical points of index p .

17.7 The basic theorems of the subject are the *Weak Morse Inequalities*: For any Morse function f ,

$$C_p(f) \geq \dim H^p(M)$$

and the *Morse index theorem*:

$$\sum_{p=0}^{\dim M} (-1)^p C_p(f) = \sum_{p=0}^{\dim M} (-1)^p \dim H^p(M).$$

17.8 In the other words, the number of critical points of index p is at least equal to the p -th Betti number. Also, the alternating sum of the number of critical points is equal to the Euler characteristic.

17.9 Examples: Any periodic function has a minimum and a maximum within a period, since S^1 has Betti numbers $B^p = 1$ for $p = 0, 1$. Any function on the torus must have at least two saddle points since $b_1(T^2) = 2$. What is the minimum number of critical points of index p of a smooth function on the Grassmannian $Gr_k(C^\infty)$?.

Chapter 18

Supersymmetry

18.1 The set of instaneous states of a q uantum mechanical system form a complex Hilbert space $\mathcal{H}$; its observables are self-adjoint operators. The spectrum of this operator is the set of outcomes when a meaurement is made of the observable. Of special importance is the hamiltonian $H : \mathcal{H} \to \mathcal{H}$, a self-adjoint operator which represents the energy of the system. Moreover, a system which is in state $\psi \in \mathcal{H}$ at time t_0 will be at time t in the state $e^{-iH(t-t_0)}\psi$.

18.2 A *conserved quantity* is an observable that commutes with the hamiltonian: $[H, Q] = 0$. It does not change under time evolution.

18.3 A system is *supersymmetric* if there are self-adjoint operators $Q_i, i = 1, \cdots N$ such that

$$Q_i Q_j + Q_j Q_i = 2\delta_{ij} H.$$

Thus, H is the square of any of the operators Q_i ; hence they are all conserved.

18.4 The case $N = 1$ is of special interest. We can write the supersymmetry algebra as

$$Q^2 = Q^{\dagger 2} = 0, \quad QQ^\dagger + Q^\dagger Q = H$$

where $Q = \frac{1}{2}(Q_1 + iQ_2)$.

18.5 The algebra generated by Q_i has an automorphism, $Q_i \to -Q_i$. There is an operator F that generates this automorphism $FQ_iF^\dagger = -F, F^2 = 1$ or equivalently

$$FQ_i + Q_iF = 0, \quad F = F^\dagger, \quad F^2 = 1.$$

The subspace of $\mathcal{H}$ with eigenvalue 1 for F is called the *bosonic* subspace and that with eigenvalue -1 the *fermionic* subspace: $\mathcal{H} = \mathcal{H}_B \oplus \mathcal{H}_F$. It is clear that $Q, Q^\dagger : \mathcal{H}_B \to \mathcal{H}_F$ and vice versa.

18.6 Within an eigenspace of $\mathcal{H}$ with non-zero eigenvalue E, the number of bosonic and fermionic states are equal; i.e, $\mathrm{tr}_E F = 0$. To see this we note that $\mathrm{tr}_E F = \frac{1}{E}\,\mathrm{tr}_E Q_1^2 F = -\frac{1}{E}\,\mathrm{tr}_E Q_1 F Q_1 = -\,\mathrm{tr}_E F$.

18.7 The states of zero energy must satisfy

$$Q_i|\psi >= 0$$

for all i. The number $W = \mathrm{tr}_0 F$ of bosonic zero energy states minus fermionic zero energy states, is called the *Witten index*. We can see that

$$W = \mathrm{tr} e^{-tH} F$$

for all $t > 0$ since only zero energy states contribute to this sum.

18.8 As an example, let $\mathcal{H}$ be the space of (complex-valued) differential forms on a real manifold M with positive Riemannian metric. The inner product is

$$< \omega, \phi >= \sum_{p=0}^{\dim M} \int \omega^*_{i_1 \cdots i_p} \phi_{j_1 \cdots j_p} g^{i_1 j_1} \cdots g^{i_p j_p} \sqrt{\det g}\, d^n x.$$

Then the bosonic subspace consists of forms of even order and fermionic states those of odd order. We choose $Q = d$, the exterior derivative.

18.9 Then the hamiltonian is the laplacian; the zero energy states are harmonic forms; i.e, forms that satisfy

$$d\omega = 0, \quad d^\dagger \omega = 0.$$

Recall that each cohomology class has a unique harmonic representative. For, if ω is closed and belongs to the orthogonal complement of the space of hamrmonic forms, then it is exact: $\omega = d(\Delta^{-1}[d^\dagger\omega])$. Thus the Witten index is the Euler characteristic of the manifold:

$$\mathrm{tr} e^{-t\Delta}(-1)^p = \sum_{p=0}^{\dim M} (-1)^p \dim H^p(M).$$

Chapter 19

Lie Groups

19.1 A *group* is a set G along with a map (*multiplication*) $G \times G \to G, (g,h) \mapsto gh$ such that: (i) it is associative, (i.e., $(gh)k = (g(hk)$) ;(ii) there is an element $1 \in G$ (the *identity*) with $g1 = 1g = g$;(iii) every element g must have an inverse g^{-1} , with $gg^{-1} = 1 = g^{-1}g$. The axioms imply that the inverse of each element is unique.

19.2 A *homomorphism* $f : G \to H$ is a map from a group to another that preserves the multiplication; $f(gh) = f(g)f(h)$. If a homomorphism is a one-one onto map it is an *isomorphism*. Isomorphic groups have exactly the same sructure.

19.3 A *Lie group* is a group which is also a manifold such that the multiplication map and the inverse map are differentiable.

19.4 The simplest example is the real line with the group operation being *addition*; or equivalently the set of positive numbers with multiplication as the group operation. The set of integers is a group but is not a manifold. The sphere S^2 is a manifold on which there exists no differentiable group operation.

19.5 There are three classical series of Lie groups, the compact simple matrix groups. These are the Lie groups of greatest imprtance in physics. The *orthogonal group* $O(n)$ is the set of all linear maps in a real n -dimensional vector space that preserve a positive inner product. (Upto isomorphism this gives a unique definition). The *Unitary group* $U(n)$ is the analogous object for complex vector spaces. It can be thought of the subgroup of $O(2n)$

that preserves a complex structure: $gJg^T = J$. And finally the *compact sympectic group* $USp(n)$ is the set of elements of $O(4n)$ that preserve a quaternionic structure; i.e., $gIg^T = I, gJg^T = J, gKg^T = K$, with $I^2 = J^2 = K^2 = -1, IJ = -JI = K$. Note that the Unitary group is a real manifold (of dimension n) although it is defined in terms of compex vector spaces.

19.6 Some subgroups of the above groups are also worth noting. $SO(n)$ is the group of orthogonal matrices of determinant one. $O(n)$ is the union of two connected components, one with determinant one and the other with determinant -1 .

Chapter 20

Poisson Brackets

20.1 The set of observables of a classical system forms a *real commutative algebra*. That is, the sum of two observables is an observable, as is its product with a real number. The product of two observables gives another observable, which is independent of the order of multiplication. Also, there is an identity element decribing the trivial observable which is always equal to one. There is one more important operation on a pair of classical observables, that of a Poisson bracket.

20.2 A *Poisson bracket* on a commutative algebra A is a bilinear map $\{,\} : \mathcal{A} \times \mathcal{A} \to \mathcal{A}$, satisfying the conditions

$$\begin{array}{ll} (i) & \{f,g\} = -\{g,f\}, \\ (ii) & \{f,\{g,h\}\} + \{g,\{h,f\}\} + \{h,\{f,g\}\} = 0 \\ (iii) & \{f,gh\} = \{f,g\}h + g\{f,h\}. \end{array}$$

A commutative algebra along with a Poisson bracket is a *Poisson algebra*.

20.3 The first two conditions say that the Poisson bracket turns the set of observables into a Lie algebra. The condition (iii) says that this bracket acts as a derivation on the commutative multiplication; i.e., satisfies the Leibnitz rule of derivation.

20.4 The complexification of a Poisson algebra can often be thought of as the limit of a one-parameter family of (complex) associative algebras, when the commutator is taken to be infinitesimally small. (Whether *every* Poisson algebra arises as such a limit is a subtle question which has been

studied by kontsevich.) The Poisson bracket is the limiting form of the commutator. The conditions on the Poisson bracket are just the remnants of the associativity of the underlying algebra. This makes sense physically if we remember that the observables of a quantum system are the (self-adjoint) elements of a complex associative algebra. The limit in which the commutator is infinitesimally small is the classical limit. We will see some examples of this idea later.

20.5 Two elements of a Poisson algebra are said to *commute* if their Poisson bracket is zero. The set of elements that commute with *all* the elements of a Poisson algebra is the *center*.

20.6 The set of instantaneous states of a classical system is called the *phase space*. A classical observable assigns a real number to each state, so it is a real valued funcion on the phase space.The sum and product of obervables correspond to the pointwise operations on these functions.

20.7 In many cases the phase space is naturally thought of as a smooth manifold. We will study this case in more detail here. An observable is then a smooth function on the phase space. (Sometimes it is more natural to think of the phase space as an algebraic variety or as as a differentiable manifold of class C^k ; the observables will then to be chosen to be the corresponding category of functions.)

20.8 Any derivation on the space of smooth functions of a manifold is a first order partial differential operator: this is the meaning of the Leibnitz rule for derivations. Explicitly in terms of co-ordinates x^i on the manifold, it takes the form

$$Vf = V^i \frac{\partial f}{\partial x^i}$$

for a set of functions V^i . In fact

$$V(x^i) = V^i.$$

These functions transform as the components of a (contra-variant) vector field under co-ordinate transformations. Thus *a derivation on the algebra of functions on a smooth manifold is the same as a vector field*. Thus a derivation on a commutative algebra is a natural generalization of the notion of a vector field.

20.9 Thus for functions on a phase space, the condition (iii) says that to every observable f there corresponds a derivation V_f :

$$V_f g = \{f, g\}.$$

This is called the *canonical* vector field of the function f .

20.10 In many mathematical texts, V_f is called the *Hamiltonian* vector field of f . However, the word Hamiltonian has a specific meaning in physics, as the generator of time translations. Therefore we will call V_f the canonical vector field of f .

20.11 In tensor notation,

$$V_f g = V_f^i \partial_i g.$$

By reversing the roles of f and g we can see that V_f^i itself is linear in the derivatives of f . Thus there must in fact be a tensor field w^{ij} such that

$$\{f, g\} = w^{ij} \partial_i f \partial_j g.$$

In fact the tensor w^{ij} gives the Poisson brackets of the co-ordinate functions with each other:

$$\{x^i, x^j\} = w^{ij}.$$

The first condition on the Poisson bracket requires that this tensor be anti-symmetric

$$w^{ij} = -w^{ji}.$$

The second condition (the Jacobi identity) then becomes

$$w^{il} \partial_l w^{jk} + w^{jl} \partial_l w^{ki} + w^{kl} \partial_l w^{ji} = 0.$$

20.12 Conversely, if we are given, on any manifold, an anti-symmetric tensor field satisfying this identity we can define a Poisson bracket on the space functions of that manifold.

20.13 An anti-symmetric tensor field w whose components satisfy the Jacobi identity (**??**) is called a *Poisson tensor* or *Poisson structure*. A manifold along with a Poisson structure is called a *Poisson manifold*.

20.14 The only Poisson bracket on the real line R is the trivial one: $\{f, g\} = 0$ for all f, g . The following is a Poisson bracket on R^2 :

$$\{f, g\} = \partial_1 f \partial_2 g - \partial_1 g \partial_2 f.$$

In this case the Poisson tensor w^{ij} has constant components,

$$w^{12} = -w^{21} = 1, \quad w^{11} = w^{22} = 0.$$

Indeed, any anti-symmetric tensor w^{ij} with constant components (in some co-ordinate system) will give a Poisson bracket on a manifold.

20.15 There are many other Poisson tensors in dimension two; see [**?**] for a classification. On three dimensional Euclidean space, we have the Poisson bracket

$$\{f, g\} = \epsilon_{ijk} x^i \partial_j f \partial_k g$$

where ϵ_{ijk} is the Levi-Civita tensor and x^i the Cartesian co-ordinate system.

20.16 There is an important difference between the two examples above: the Poisson tensor w^{ij} is non-degenerate (has non-zero determinant) in the case of R^2 but it clearly vanishes at the origin in the case of R^3 . Even away from the origin, the tensor $w^{ij} = \epsilon_{ijk} x^k$ has a zero eigenvector:

$$w^{ij} x^j = 0.$$

As a result, a function on R^3 that depends only on the radial distance will have zero Poisson bracket with any function: its derivative points in the radial direction. This leads us to a concept of non-degeneracy.

20.17 A Poisson structure is *non-degenerate* if

$$\{f, g\} = 0 \text{ for all } g \Rightarrow f = \text{ constant.}$$

In other words a Poisson structure is non-degenerate if the corresponding Poisson algebra has a trivial center; i.e., consists of just the multiples of the identity.

20.18 A Poisson structure is non-degenerate if and only if the matrix w^{ij} has non-zero determinant everywhere. Thus, non-degenerate Poisson brackets exist only on even-dimensional manifolds: recall that the determinant of any anti-symmetric matrix of odd dimension is zero.

20.19 On R^{2n}, suppose we denote the first n co-ordinates by q^a (for $a = 1, 2, \cdots n$) and the next n by p_a. The following brackets define a non-degenerate Poisson structure on R^{2n}:

$$\{q^a, q^b\} = \{p_a, p_b\} = 0, \quad \{p_a, q^b\} = \delta^b_a.$$

This is called the *standard* Poisson structure on R^{2n}. The Poisson tensor has components $w^{ij} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ in this co-ordinate system. A co-ordinate system in which the above Poisson brackets hold is called a *canonical* co-ordinate system.

20.20 If a Poisson-tensor is non-degenerate, there is always a co-ordinate system that makes it a constant within each co-ordinate neighborhood. The Jacobi identity is just the integrability condition that makes this possible. We can even choose the co-ordinate system so that $w^{ij} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. Thus locally every *non-degerate* Poisson structure is equivalent (upto co-ordinate transformations) to the standard Poisson structure on R^{2n}. This result is known as 'Darboux theorem'. Thus there is no local geometric information in a symplectic structure: there is no symplectic analogue of the curvature of Riemannian geometry.

20.21 Let $\underline{G}$ be a Lie algebra with basis e^i and commutation relations

$$[e^i, e^j] = c^{ij}_k e^k.$$

The structure constants are then anti-symmetric

$$c^{ij}_k = -c^{ji}_k$$

and satisfy the Jacobi identity

$$c^{ij}_l c^{lk}_m + c^{jk}_l c^{li}_m + c^{ki}_l c^{lj}_m = 0.$$

Consider the dual vector space $\underline{G^*}$ of $\underline{G}$ and let ξ^i be the co-ordinate defined on it by the basis e^i of $\underline{G}$. Now we can define a Poisson bracket on $\underline{G^*}$ by

$$\{\xi^i, \xi^j\} = c^{ij}_k \xi^k.$$

The Poisson structure $w^{ij} = c^{ij}_k \xi^k$ satisfies the Jacobi identity due to the identity on the structure constants of a Lie algebra.

20.22 If we regard R^3 as the (dual of the) Lie algebra of the rotation group, its structure constants are just the components of the Levi-Civita tensor. Thus our example of a Poisson structure on R^3 is just a special case of the last example.

20.23 A vector field can be thought of as an infinitesimal transformation of a manifold to itself. To visualize this, think of the points of the manifold as though they are particles in a fluid, with the vector field having the meaning of the velocity of the fluid. Suppose the components of a vector field are $V^i(x)$ in some co–ordinate system. An infinitesimal 'time' t later the point x^i will be mapped to the nearby point $x^i + V^i(x)dt$. By repeating this process we can obtain a one-parameter family of transformations ϕ_t of the manifold to itself. They are solutions to the differential equations

$$\frac{d\phi^i_t(x)}{dt} = V^i(\phi_t(x)),$$

with the initial condition that at 'time' $t = 0$ we get the identity map: $\phi^i_0(x) = x^i$. Moreover, these transformations satisfy the composition law

$$\phi_{t_1}(\phi_{t_2}) = \phi_{t_1+t_2}$$

which means that they form a one-parameter group. More precisely, they define an action of the additive group of real numbers on the manifold.

20.24 Now, each function on a Poisson manifold defines a vector field V_f with components $V^i_f = w^{ij}\partial_j f$. Thus, associated to each observable on the phase space of a classical mechanical systems is a one-parameter group of transformations of the manifold to itself. These are the *canonical transformations* generated by f. Of course any vector field generates such a one parameter group of transformations; the canonical transformations are special in that they leave the Poisson structure invariant. In a sense they are the "inner automorphisms" of the Poisson structure.

20.25 Now we can see another interpretation of the Poisson bracket:

$$\{f, g\} = V_f^i \partial_i g$$

is just the infinitesimal change of g under the canonical transformation generated by f .

20.26 Time evolution of a classical mechanical system is a one-parameter group of canonical transformations; the observable that generates this transformation is the *Hamiltonian*. Thus the Hamiltonian is the single most important observable of a mechanical system. It has the physical meaning of energy.

20.27 A free particle of mass m moving in R^3 has a six-dimensional phase space: its instantaneous state is described by the position co-ordinates q^a , for $a = 1, 2, 3$ and the momentum variables p_a . The Poisson structure is the standard one:

$$\{q^a, q^b\} = 0 = \{p_a, p_b\}, \quad \{p_a, q^b\} = \delta_a^b.$$

The Hamiltonian is just the kinetic energy,

$$H = \frac{p_a p_a}{2m}.$$

Thus we have

$$\{H, q^a\} = \frac{1}{m} p^a, \quad \{H, p_a\} = 0$$

The canonical vector field of H is

$$V_H = \frac{1}{m} p^a \frac{\partial}{\partial q^a}.$$

Suppose the canonical transformations generated by the Hamiltonian map the point (q, p) to $(q(t), p(t))$ at time t . Then,

$$q^a(t) = q^a + \frac{1}{m} p^a t, \quad p_a(t) = p_a.$$

This describes the motion of a free particle in a straight line. Thus the momentum is a constant under time evolution and $\frac{p_a}{m}$ has the meaning of velocity.

20.28 An observable whose Poisson bracket with the Hamiltonian vanishes is a *constant of the motion*: it does not change its value under time evolution.

20.29 In the above example, momentum is also a constant of the motion. The Hamiltonian itself is always a constant of the motion; its conservation is just the law of conservation of energy. (We will only consider 'closed' systems for which the Hamltonian and the constraints have no explicit dependence on time.) The Jacobi identity can be used to show that if f and g are constants of the motion, so is $\{f, g\}$. Thus the set of all constants of the motion form a Lie algebra under the Poisson bracket. Since the product of two constants of motion is also one, the set of constants of motion form a sub-Poisson algebra.

20.30 If a Poisson tensor w is non-degenerate (i.e., the Poisson algebra has trivial center), we can find a two-form ω which is the inverse of w :

$$\omega = \omega_{ij} dx^i \wedge dx^j, \quad \omega_{ij} w^{jk} = \delta_i^k.$$

The Jacobi identity on w becomes a much simpler, *linear*, equation in terms of ω :

$$\partial_i \omega_{jk} + \partial_j \omega_{ki} + \partial_k \omega_{ij} = 0.$$

In other words, ω is a closed two-form. Conversely any closed non-degenerate two-form defines a Poisson tensor.

20.31 A *sympletic form* or *symplectic structure* is a closed nondegenerate two-form. A manifold along with a sympletic structure is a symplectic manifold.

20.32 A symplectic manifold is even-dimensional.

20.33 The standard Poisson structure on $R^{2n} = \{(q^i, p_i), i, j = 1, \cdots n$ arises from the sympletic form

$$\omega = dq^i \wedge dp_i.$$

20.34 The volume form is a sympletic structure on the sphere S^2 .

20.35 The rigid body is a system which has led to much beautiful mathematics and physics [**?**]. If the distance between any two points of a body is fixed as it moves, it is a rigid body. The group of transformations on space that preserve the distance between any pair of points is its isometry group. In the case of R^3 this is the semi–direct product of rotations and translations $O(3) \times R^3$. Any configuration of a rigid body can be mapped into any other by an action of the isometry group of space. [1] Since time evolution is continuous, we can restrict to the connected component of the isometry group. Also, it is convenient to consider the center of mass to be at rest, so that we can restrict ourselves to configurations that can be mapped into each other by rotations. We will study the rigid body in the absence of external torques.

20.36 Thus we can describe the time evolution of a rigid body by a curve $g : R \to SO(3)$ in the rotation group: the configuration at time time t is obtained by a rotation $g(t)$ from some arbitrarily chosen reference configuration. We will use the convention that time evolution is a left action by $g(t)$ on this initial point in the configuration space. The angular velocity of the body relative to space (i.e., w.r.t. an observer in the same reference frame as the center of mass) is then the generator of left translations, hence a *right* invariant vector field of the group $SO(3)$. Explicitly, the angular velocity relative to space is

$$\omega_s = \frac{dg}{dt} g^{-1}$$

which can be thought of as an element of the Lie algebra $SO(3)$. The angular velocity in a reference frame moving with the body differs from this by a rotation (adjoint action) by $g(t)$:

$$\omega_c = g^{-1} \frac{dg}{dt}.$$

20.37 The kinetic energy of the system is given by a positive quadratic form A on the Lie algebra,

$$H = \frac{1}{2} A(\omega_c, \omega_c).$$

[1] For a rigid body of generic shape, this action will not have fixed points: any rotation or translations will change the configuration

A is the moment of inertia. (See [**?**] for a derivation from first principles.) The quantity $L = A(g^{-1}\frac{dg}{dt})$ is the angular momentum relative to the body. The angular momentum relative to space differs from this by a rotation by $g(t)$,

$$R = gLg^{-1}.$$

In the absence of external torques, this is a conserved quantity. The conservation of R leads to the equations of motion:

$$\frac{dL}{dt} = [L, g^{-1}\dot{g}].$$

These are the equations of motion of a rigid body.

20.38 We can solve for the angular velocity to express the equations of motion in terms of the angular momentum. It is convenient to do this in a basis that diagonalizes the moment of inertia:

$$A = \begin{pmatrix} a_1 & 0 & 0 \\ 0 & a_2 & 0 \\ 0 & 0 & a_3 \end{pmatrix}.$$

We will get,

$$\frac{dL_1}{dt} = -b_1 L_2 L_3, \quad \frac{dL_2}{dt} = b_2 L_3 L_1, \quad \frac{dL_3}{dt} = -b_3 L_1 L_2.$$

These are the famous [**?**] Euler equations of a rigid body. The constants here are defined by

$$b_1 = a_2^{-1} - a_3^{-1}, \quad b_2 = a_1^{-1} - a_3^{-1}, \quad b_3 = a_1^{-1} - a_2^{-1}.$$

20.39 If the constants a_1, a_2, a_3 are all equal, the Euler equations are trivial to solve: each component of L_a is a constant of the motion. If even a pair of the constants are equal, (say $a_1 = a_2$) then L_3 is conserved and the solution in terms of trigonometric functions is elementary. So we will consider the non-isotropic case from now on. We can assume without any loss of generality that $a_1 < a_2 < a_3$; then b_1, b_2, b_3 are positive.

20.40 The equations for L_a do not involve g ; so it is reasonable to solve them first, forgetting about the variables g for the moment. The phase space is now just R^3 , with the components of angular momentum L_i forming the Cartesian co-ordinate system . The hamiltonian can be expressed as a quadratic function:

$$H = \frac{L_1^2}{2a_1} + \frac{L_2^2}{2a_2} + \frac{L_3^2}{2a_3}.$$

If we postulate on R^3 the Poisson brackets of the earlier example,

$$\{L_i, L_j\} = \epsilon_{ijk} L_k$$

and calculate the Poisson brackets of the angular momentum with the Hamiltonian, we get the Euler equations of motion.

20.41 We have two obvious constants of the motion: the hamiltonian H and the central function

$$L^2 = L_1^2 + L_2^2 + L_3^2.$$

In the case of a non–isotropic rigid body these constant of motion are independent of each other. The surfaces of fixed energy are ellipsoids; the surfaces of fixed L^2 are spheres. Thus the time evolution of the system must take place on the intersection of these two surfaces. Thus it shouldn't be surprising that the Euler equations can be solved explicitly in terms of elliptic functions.

20.42 The Jacobi elliptic functions [**?**] of modulus k satisfy the quadratic constraints

$$\mathrm{sn}^{\,2}(u,k) + \mathrm{cn}^{\,2}(u,k) = 1, \quad k^2\,\mathrm{sn}^{\,2}(u,k) + \mathrm{dn}^{\,2}(u,k) = 1$$

and the differential equations

$$\begin{aligned}
\frac{d\,\mathrm{cn}\,(u,k)}{du} &= -\,\mathrm{sn}\,(u,k)\,\mathrm{dn}\,(u,k), \\
\frac{d\,\mathrm{sn}\,(u,k)}{du} &= \mathrm{cn}\,(u,k)\,\mathrm{dn}\,(u,k), \\
\frac{d\,\mathrm{dn}\,(u,k)}{du} &= -k^2\,\mathrm{sn}\,(u,k)\,\mathrm{cn}\,(u,k).
\end{aligned}$$

It is not a coincidence that these differential equations look analogous to the Euler equations of the rigid body; Jacobi [**?**] discovered these functions by solving the motion of the rigid body.

20.43 The ansatz

$$L_1 = A_1 \,\mathrm{cn}\,(\omega t, k), \quad L_2 = A_2 \,\mathrm{sn}\,(\omega t, k), \quad L_3 = A_3 \,\mathrm{dn}\,(\omega t, k)$$

will then solve the Euler equations. More precisely the differential equations are reduced to algebraic relations for the constants:

$$\omega A_1 = b_1 A_2 A_3, \quad \omega A_2 = b_2 A_1 A_3, \quad k^2 \omega A_3 = b_3 A_1 A_2.$$

We get, upon solving these,

$$A_1^2 = \frac{1}{b_2}(2H - \frac{L^2}{a_3}), \quad A_2^2 = \frac{1}{b_1}(2H - \frac{L^2}{a_3}), \quad A_3^2 = \frac{1}{b_2}(\frac{L^2}{a_1} - 2H),$$

and

$$\omega^2 = b_2(\frac{L^2}{a_1} - 2H), \quad k^2 = \frac{a_1 b_3 (2a_3 H - L^2)}{a_3 \sqrt{(b_1 b_2)(L^2 - 2a_1 H)}}.$$

Chapter 21

Deformation Quantization

21.1 The set of instantaneous states of a quantum system form a complex Hilbert space $\mathcal{H}$; the algebra of linear operators is an associative algebra; the self-adjoint operators are the *observables*. Note that the set of observables is not an algebra, since the product of two observables is not always an observable, because it may not be self-adjoint.

21.2 In the limit as $\hbar \to 0$, the observables must form a Poisson algebra. To see how this happens let us start with a system whose phase space is $R^{2n} = \{(x, p)\}$ with the standard Poisson bracket

$$\{A, B\} = \frac{\partial A}{\partial x^i}\frac{\partial B}{\partial p_i} - \frac{\partial A}{\partial p_i}\frac{\partial B}{\partial x^i}$$

21.3 The usual rules of quantum mechanics (Schrodinger quantization) say that the quantum Hilbert space $\mathcal{H} = L^2(R^n)$. An operator $\hat{A}$ in this space can always be represented by an integral kernel

$$\hat{A}\psi(x) = \int \tilde{A}(x, y)\psi(y)dy.$$

Self-adjoint operators have hermitean kernels: $\tilde{A}(x, y) = \tilde{A}^*(y, x)$. The product of operators cooresponds to a multliplication of the integral kernels

$$\widetilde{AB}(x, y) = \int \tilde{A}(x, z)\tilde{B}(z, y)dz.$$

21.4 We expect that corresponding to each operator $\hat{A}$ there is a function $A(x,p)$ in the classical phase space. Given an operator with integral kernel $\tilde{A}(x,y)$ we define its *Weyl symbol*

$$A(x,p) = \int \tilde{A}(x+\frac{z}{2}, x-\frac{z}{2})e^{-\frac{i}{\hbar}p\cdot z}dz.$$

This is a partial Fourier transform of the kernel, where the relative coordinate is replaced by its Fourier conjugate. The inverse map associates to every function on the phase space a linear operator with integral kernel) *correspondence principle*)

$$\tilde{A}(x,y) = \int A(\frac{x+y}{2}, p)e^{\frac{i}{\hbar}p\cdot(x-y)}[dp].$$

Here $2\pi\hbar$ is Planck's constant and $[dp]$ stands for $\frac{d^n p}{(2\pi\hbar)^n}$.

21.5 This map from functions on the phase space and operators on the Hilbert space has all the properties we expect of the classical to quantum correspondence. Real functions go to hermitean integral kernels and vice versa. The trace of an operator is the integral of its symbols:

$$\mathrm{tr}\hat{A} = \int \tilde{A}(x,x)dx = \int A(x,p)dx[dp].$$

21.6 A function independent of momenta correspond to a multiplication operator:

$$\hat{A}\psi(x) = A(x)\psi(x)$$

as we expect in the Schrodinger picture of quantum mechanics.

21.7 The operator corresponding to p_i is the derivative operator : $\hat{p}_k = \frac{\hbar}{i}\frac{\partial}{\partial x^k}$. The simplest way to see this to consider the function $A(x,p) = e^{ip_i a^i}$ whose operator kernel is

$$\tilde{A}(x,y) = \delta(x+\hbar a - y)$$

so that the operator $\hat{A} = e^{\hbar a^k \partial_k}$ is the translation by $\hbar a$. For infinitesimally small a we get the derivative operator.

21.8 A polynomial in p with coeffiecients which are functions of x corresponds to the operator in which p is replaced by the derivative operator, with the *symmetric* ordering rule:

$$f(x)p^{l_1}\cdots p^{l_b} \to \frac{1}{b+1}\sum_{a=0}^{b}\hat{p}^{l_1}\cdots\hat{p}^{l_a}f(x)\hat{p}^{l_{a+1}}\cdots\hat{p}^{l_b}.$$

For example,

$$x^k p_l \to -i\hbar\frac{1}{2}[x^k\partial_l + \partial_l x^k].$$

Thus we recover the usual form of the correspondence principle used in elementary quantum mechanics.

21.9 It is important that the Weyl symbol is not just the classical approximation to the quantum operator. It contains *all* the information in the operator, indeed the operator is determined by its symbol by the above Fourier transformation. Thus it should be possible to translate the multiplication of operators to the classical language: $A\circ B$ is the symbol of the operator $\hat{A}\hat{B}$. it must correspond to some multiplication law for function in phase space.

21.10 The above formulae yield the following result:

$$A\circ B(x,p) = \left\{ e^{\frac{-i\hbar}{2}(\frac{\partial}{\partial x^i}\frac{\partial}{\partial p'_i}-\frac{\partial}{\partial p_i}\frac{\partial}{\partial x i'})} A(x,p)B(x',p')\right\}_{x=x';p=p'}.$$

To understand this better we expand in powers of $\hbar$:

$$\tilde{A}\circ\tilde{B}(x,p) = \sum_{n=0}^{\infty}\left(\frac{-i\hbar}{2}\right)^n\frac{1}{n!}\{\tilde{A},\tilde{B}\}_{(n)}.$$

where

$$\{\tilde{A},\tilde{B}\}_{(n)} = \sum_{r=0}^{n}(-1)^r\tilde{A}^{j_1\cdots j_r}_{i_1\cdots i_{n-r}}\tilde{B}^{i_1\cdots i_{n-r}}_{j_1\cdots j_r}$$

where $\tilde{A}^i = \frac{\partial\tilde{A}}{\partial p_i}$ and $\tilde{A}_i = \frac{\partial\tilde{A}}{\partial x^i}$, $A^{j_1\cdots j_s}_{i_1\cdots i_r} = \frac{\partial^{r+s}A}{\partial p_{j_1}\cdots\partial p_{j_s}\partial x^{i_1}\cdots\partial x^{i_r}}$ etc. It is possible to prove by induction on the order that this multiplication is associative.

21.11 The first order term is just the Poisson bracket:

$$A \circ B(x,p) = A(x,p)B(x,p) - \frac{i\hbar}{2}\{A,B\} + O(\hbar^2)$$

Thus, *the Poisson bracket is the first order deviation way from the pointwise product of observable.* The Poisson bracket is a remnant in the classical theory of the non-commutativity of quantum mechanics.

21.12 Thus quantization is nothing but a change (*deformation*) in the multiplication law for functions on phase space, to a new law that is associative but not commutative. The product involves the value of the function at nearby points as well: it is nonlocal in phase space. This is a consequence of the uncertainty principle, the points in phase space no longer correspond to states in the quantum theory.

21.13 The deformation must be such that to first order it reproduces the Poisson bracket. But this does not fix it uniquely: there may be more than one quantum theory corresponding to a given classical theory. We have given the rules that give the usual Schrodinger quantization. Such ambiguities are present in any quantization scheme.

21.14 Deformation quantization gives a convenient starting point for deriving the semi-classical approximation. A typical question of interest in the quantum mechanics is the determination of the spectrum of an operator $\hat{h}$: i.e., the set of singularities of the resolvent operator $\hat{R}(E) = (\hat{h} - E)^{-1}$. The symbol of the resolvent operator satisfies

$$R(E) \circ (h - E) = 1.$$

21.15 This can be solved as a power series in $\hbar$:

$$R(E) = \sum_{k=0}^{\infty} R_{(k)}(E)\hbar^k,$$

$$\sum_{n=0}^{\infty}\sum_{k=0}^{\infty}\left(\frac{-i\hbar}{2}\right)^n \frac{1}{n!}\hbar^k\{R_{(k)}(E), h-E\}_{(n)} = 1.$$

Equating the powers of $\hbar$ on both sides of this equation, we get a set of recursion relations

$$R_{(0)}(E) = (h - E)^{-1},$$

$$R_{(m)}(E) = -\sum_{n=1}^{m} (\frac{-i}{2})^n \frac{1}{n!} \{R_{(n-m)}(E), h\}_{(n)} (h - E)^{-1}.$$

21.16 Of particular interest is the case where h is of the form:

$$h(x, p) = t(p) + v(x).$$

In this case the mixed derivatives in the generalized Poisson brackets vanish and we get

$$\{R_{(k)}, h\}_{(n)} = R_{(k)i_1 \cdots i_n} t^{i_1 \cdots i_n} + (-1)^n \tilde{R}^{i_1 \cdots i_n}_{(k)} v_{i_1 \cdots i_n}.$$

If $t(p) = p_i p_i$ as in nonrelativistic quantum mechanics, there is only one term for $n > 2$,

$$\{R_{(k)}, h\}_{(n)} = (-1)^n R^{i_1 \cdots i_n}_{(k)} v_{i_1 \cdots i_n}$$

while

$$\{R_{(k)}, h\}_{(1)} = 2p_i R_{(k)i} - R^i_{(k)} v_i$$

and

$$\{R_{(k)}, h\}_{(2)} = 2R_{(k)ii} + R^{ij}_{(k)} v_{ij}.$$

In fact, $R_{(1)} = 0$ and

$$R_{(2)}(E) = \frac{1}{2(h - E)^2} [\frac{v_i v_i}{h - E} + 2(p_i v_i)^2 - v_i i].$$

These formalae are useful for example in deriving the Thomas-Fermi theory in atomic physics.

21.17 In general the observables of a classical mechanical system form a Poisson algebra. Kontsevich has found a way of deforming such a general Poisson algebra into an associative algebra, a generalization of the above infinite series. The physical applications are still being worked out.